

[논문]

## 인지과학 방법론 연구\*

김 영 정\*\*

**【주제분류】** 인지과학, 과학철학, 심리철학, 방법론

**【주요어】** 인지과학 방법론, 환원, 기능적 분석, 병렬분산처리, 분야간 이론

**【요약문】** 이 논문은 인지과학의 방법론을 개괄적으로 살펴보는 것을 목적으로 한다. 먼저 상위수준의 인지과학과 하위수준의 신경과학의 이론들을 연결시키는 방법론으로 환원, 기능적 분석, 병렬분산처리 등의 수준간 연계 방법론을 차례로 살펴본 후, 동일하게 상위 수준에 속하는 심리학, 인공지능, 언어학, 철학 등과 같은 인지과학내 학문 분야간의 이론들을 연결시키는 방법론으로 '경계 허물기'와 '경계에 다리놓기' 등의 수준내 연계 방법론을 살펴본다.

이를 토대로 앞서 살펴본 '환원', '기능적 분석', '병렬분산처리', '경계 허물기', '경계에 다리놓기' 등을 모두 포괄하는 방법론으로 '분야간 이론(interfield theory)'을 제안한다. 이 분야간 이론은 위의 학문 분야의 이론들을 모두 하나의 틀에 맞춰 통일적으로 설명해내는 그러한 획일적인 방법론이 아니라 다양한 형태의 틀들을 아우르는 방식의 방법론이다. 이 분야간 이론은 다른 학문 분야들에서 각기 다른 지향점을 가지고 연구된 현상들간의 연결을 끊어내고 그들간의 관계에 대한 전체적인 조망을 가능케 해준다는 점에서 포괄적 방법론으로 간주될 수 있다. 비유하여 말하자면, 인지과학 방법론은 적어도 지금 시점에서는 모든 다양한 종류의 철광석들을 하나의 방식으로 녹여내는 용광로(melting pot)일 수는 없으며, 다양한 재료들을 모아 그 전체성에서 조화를 이루어내는 샐러드 보울(salad bowl)이 될 수밖에 없다는 것이 필자의 결론이다. 인지과학 연구는 다원적 형태를 아우르는 방식의 방법론으로 진행되어야 할 것이다.

\* 이 연구는 1999년도 서울대학교 발전기금 일반 학술연구비 지원에 의해 수행되었음.

\*\* 서울대 철학과 교수

지난 몇 십 년 동안 철학계에서 일어난 가장 눈에 띄는 특징은 아마도 과학적 탐구에 대한 논의가 갈수록 늘어나는 현상일 것이다. 이런 특징은 마음에 대한 철학적 탐구인 심리철학 분야에서는 더 한층 심하다. 심리철학자들은 심리학과 신경과학의 연구 결과들을 관심 있게 바라본다. 생각, 지각, 추론 같은 정신 현상을 철학적으로 이해하기 위해 인지에 대한 생물학적 모형들이 판단 근거가 되고 있다. 그렇지만 이런 일들은 그렇게 놀랄 일은 아니다. 우리는 철학사를 조금만 들여보아도 철학 속에 그러한 경험적 연구들이 자리잡고 있음을 알 수 있다. 근세의 경험론자들은 관념의 이론을 가지고 인지에 대한 이론적인 모형을 개발하고 발전시켰다. 비록 지금의 관점에서 보면 서투른 것이긴 했지만, 그 당시로서는 분명 경험적인 탐구였다. 요즘의 인식론자들은 그런 경험론자들의 후예다. 그들이 지식에 관심이 있고, 추론을 지배하는 규범을 찾으려고 한다면, 정신 표상과 추론 구조에 대해 지금 경험과학적으로 가장 뛰어난 지식, 곧 심리학, 신경과학 등에서 무엇인가를 얻으려고 기대하는 것은 전혀 이상한 일이 아닐 것이다.

철학자들이 마음에 대한 경험 과학에 갈수록 관심을 넓혀 가는 데에는 학문 사이의 칸막이가 무너진 데 그 배경이 있다. 인공지능 분야에서 연구하는 사람들은 인간의 여러 가지 인지 능력을 모사하려고 노력한다. 심리학자들은 가설을 세우고 검증하는 데 컴퓨터 모의실험을 사용하고 있다. 이렇게 인지와 마음에 대해 여러 학문들이 함께 손잡고 연구하는 경향에서 인지과학(cognitive science)은 태어나게 된 것이다.

인간의 마음과 컴퓨터의 인지 현상에 대한 탐구를 목표로 하는 인지과학은 1950년대를 기점으로 태동을 시작하여 1970년대 중반에 비교적 뚜렷한 모습을 갖추어 부상한 다학문적 종합과학이다. 현재 인지과학에는 심리학, 인공지능학, 신경과학, 언어학, 인류학, 철학 등이 참여하여, 활발히 공동 연구가 진행되고 있다. 인지과학의 태동기에 여러 학문들을 '인지과학'이라는 이름표 아래 묶어주는 것은 계산적 정보처리 모형의 방법과 개념을 사용한다는 점이었고, 이 모형은 아직도

인지과학에서 영향력 있는 한 방법론으로 자리잡고 있다. 컴퓨터뿐만 아니라 인간의 마음도 계산적 정보 처리 체계라는 아이디어가 인지과학 탄생에 토대 역할을 하였던 것이다. 인지과학의 한 방을 차지하는 철학에서도 마찬가지이다. 컴퓨터가 생겨나 정신 현상에 대해 새로운 시각에서 접근해 볼 수 있게 되었고, 전통적인 심신 문제(mind-body problem)도 새롭게 궁리해 보게 되었다. 가령 심리철학의 기능주의는 마음이 곧 튜링 기계라고 주장함으로써 뜨게 된 이론이다. 이렇게 인간의 마음이 정보 처리 체계라는 아이디어를 주장하는 인지과학의 학설을 보통 계산적 기능주의 또는 정보 처리 기능주의라고 부른다.

인지과학의 주된 연구는 방금 말했듯이 마음을 계산적 정보 처리 체계로 생각함으로써 착수된다. 그러나 계산적 정보 처리 체계 접근 방식이 완전히 인지과학의 방법론을 결정하는 것은 아니다. 대부분의 인지과학이 여기에서 논의하고 있는 정보 처리 모형에 잘 들어맞기는 하지만, 마음에 대한 연구에 훨씬 더 신경 과학을 토대로 한 접근 방식을 내세우는 사람들도 있다. 이들은 상위과학이 하위과학으로 환원된다는 환원주의를 전제로 출발하는데, 환원주의는 특히 마음에 대해서 상위과학인 심리학이 하위과학인 신경과학으로 환원된다고 주장한다. 계산적 정보 처리 접근 방식은 인공지능과 인지심리학에서 채택되는 것으로 '하향식(top-down)' 방식이며, 신경과학에서 채택하고 있는 '상향식(bottom-up)' 방식 또한 인지과학의 한 주요 방법론이다.

이 논문은 인지과학의 방법론을 개괄적으로 살펴보는 것을 목적으로 한다. 인간과 컴퓨터의 인지 현상을 대상으로 하는 인지과학은 여러 학문의 공동연구를 필요로 하기 때문에, 인지과학에 관여하는 제반 학문 분야를 아우를 수 있는 방법론적 틀의 확립이 필요하다. 그러나 그러한 포괄적 방법론적 틀은 아직 확립되어 있지 않은 상태이다. 이런 사정은 인지과학의 역사가 짧은 까닭에 국내뿐만 아니라 구미에서도 마찬가지이다. 이 논문은 위에서 언급한 하향식 방식과 상향식 방식의 방법론을 중심으로 인지과학 방법론을 전체적으로 조망해보는 데 그 목적이 있다.

## I. 수준간 방법론으로서의 환원

우리가 검토할 첫 번째 방법론은 심리학이 신경과학으로 환원된다고 주장하는 환원주의(reductionism), 곧 상향식 방법론이다.

환원의 개념을 엄밀히 함으로써 환원주의를 자신들의 명시적 프로그램으로 옹호한 최초의 이들은 20세기 초반의 논리실증주의자들이었다. 논리실증주의의 환원 개념과 환원주의 프로그램은 이후 많은 비판을 받고 쇠퇴의 길을 걸어갔다. 그러나 환원과 관련된 문제들은 논리실증주의의 뒤를 이은 논리경험주의에 의해 여전히 중요한 철학적 문제들로 간주되었고, 지속적으로 논의되어 왔다.

환원주의에서 말하는 환원의 가장 기본적인 대상은 과학 이론이다. 왜냐하면 이론 이외에 우리가 흔히 환원의 대상으로 여기는 속성, 사건, 현상과 같은 것들은 모두 이론들에서 가정되는 것들이며, 그러므로 그것들 사이의 환원 관계는 그것들과 관련된 이론들의 환원 관계에서 파생적으로 따라 나온다고 간주할 수 있기 때문이다. 가령 온도라는 현상이 분자의 평균 운동 에너지로 환원된다고 우리가 말할 수 있는 근거는 온도와 관련된 이론인 열역학이 분자의 운동 에너지와 관련된 이론인 통계역학으로 환원될 수 있기 때문이다. 마찬가지로 인지 상태나 정신 속성이 신경의 상태나 속성으로 환원될 수 있는지 여부는, 인지 상태나 정신 속성의 본성과 관련된 이론인 심리학이 신경의 상태나 속성과 관련된 이론인 신경과학으로 환원될 수 있는지의 결과 여부에 따라 파생적으로 결정될 수 있다.

### 1. 환원의 장점들

위와 같은 이론간의 환원은 두 이론 사이에 설명적인 통합(explanatory unification)을 가져온다. 이를 통해서 우리는 다양한 분

야의 이론들간의 관계를 정립할 수 있고, 어떤 이론이 기술하는 현상을 더욱 잘 이해할 수 있다. 가령 위에서 예를 든 열역학과 통계역학의 경우를 보면 우리는 둘 사이의 환원 관계가 성립한다는 것을 앞으로 전자가 다루는 현상을 후자를 통해 설명할 수 있고, 이에 따라 열역학에서 가정하는 법칙들이 왜 성립하는지를, 그리고 열역학이 서술하는 온도와 같은 현상의 본성이 무엇인지를 더욱 잘 이해하게 된다. 마찬가지로 인지과학 내에서도 심리학이 신경과학으로 환원된다는 것을 앞으로 인지 상태나 정신 속성의 본성을 신경의 상태나 속성을 통해서 이해할 수 있게 된다.

다른 한편으로 이론들간의 환원은 두 이론이 각기 달리 수용하고 있는 속성이나 현상을 동일시할 수 있는 근거를 마련해줌으로써 존재론적 단순성(ontological simplification)을 가져온다. 이론간의 환원을 통해, 온도를 기체의 평균 분자운동 에너지와, 물을  $H_2O$ 와, 번개를 전기의 방전과 동일시할 수 있는 것처럼, 우리는 이전에 두 이론이 각각 자신의 법칙들을 통해서 서로 다른 두 가지 종류의 현상으로 특징지어왔던 것들이 실제로는 한 가지 종류의 현상에 지나지 않는다는 것을 알게 됨으로써 쓸데없이 존재자의 수를 늘리지 않게 된다. 따라서 심리학이 신경과학으로 환원된다면 인지 현상은 다른 아닌 신경의 작용과 같다는 것을 우리는 알게 되는 것이다.

## 2. 환원에 대한 여러 비판들

환원이 가지는 설명적 통합이나 존재론적 단순성과 같은 강점에도 불구하고, 많은 과학자들은 심리학의 이론들을 신경과학의 이론들로 환원하는 것이 가능하지 않음을 보이고자 하였다. 포더(1974)는 심리학 이론의 용어들과 신경과학 이론의 용어들을 동등화하는 교량 법칙들을 구축할 수 없음을 주장함으로써 이러한 기획을 시도하였다. 그는 우리가 두 이론의 용어들을 동일한 자연 현상을 지칭하는 데에 사용함에도 불구하고 그것들을 동일한 것으로 간주할 수 없다고 주

장한다. 심리학적 용어(예를 들어 휴가를 계획하는 것)와 신경과학적 용어(신경 흥분의 양태를 기술하는 것)가 어떤 경우에는 두뇌의 동일한 상태를 기술할 수 있을지라도 또 다른 경우들에 있어서는 이 기술들이 다른 상태들을 지칭할 수 있다는 것이다. 그 까닭은 심적 현상을 위해 사용된 분류의 도식(classification scheme)은 신경생리학적 현상을 기술하기 위하여 사용된 것과는 다른 목적을 위해 차용된 것이어서 두 어휘체계들이 사물들을 서로 다른 방식으로 분류하기 때문이다. 한 학문분야의 어휘체계는 우리가 그 분야에서 표현하고자 하는 관계들을 언명하기 위하여 설정되기 때문에, 심리학은 신경과학에서 이용하는 것과는 다른 방식으로 사건들을 분류할 이유를 갖는다. 그러므로 포더의 주장에 따르면 심리학과 신경과학은 서로 다른 종류의 관계들을 포착하고자 하고 있으며, 이렇게 되면 불가통약적인 분류체계들(incommensurable classification systems)이 생겨난다. 만약 심리학적 용어체계와 신경과학적 용어체계 사이의 분류 방식이 서로 다르다면, 두 학문 분야의 어휘들간의 교량 법칙은 있을 수 없을 것이며 따라서 양자의 법칙들을 서로에게로 환원하는 일도 할 수 없을 것이다.

심리학적 용어들과 신경 과학의 용어들을 동등화할 수 없다고 생각하는 것에 대해 흔히 언급되는 또 다른 이유는 복수실현가능성(multiple realizability)에 그 기반을 두고 있다. 우리는 인간 이외의 다른 유기체나 컴퓨터와 같은 대상들에게서는 인간과 동일한 신경과정들이 발생하지 않음을 알고 있음에도 불구하고 그것들의 활동을 기술할 때 인간의 활동을 기술하는 심리학적 용어들을 그대로 사용한다는 것이다(퍼트남 1975a). 다시 말해, 동일한 심리학적 현상이 인간의 두뇌와 동물의 두뇌뿐만 아니라 컴퓨터나 가능적으로는 외계인의 두뇌와 같은 다른 기체들(substrates)에서도 실현될 수 있기 때문에, 심리학적 현상들의 분류와 신경과학적 현상들의 분류를 동등화하는 것은 불가능하다는 것이다.

더 나아가서 포더(1978)는 우리가 심리학적 용어들을 신경 용어들

과 일치시킬 경우 심리학의 설명 능력의 상당부분을 포기하여야 할 것이라고 주장한다. 포더의 관점에 따르면 심리학의 과제 중 일부는 이성적인 인간의 행위를 설명하는 것이다. 이것을 위해서 우리는 명제 태도(즉 믿음)의 용어들로 인간의 심리학적 사태를 기술할 수 있어야 한다. 명제의 내적 구조는 우리의 심리학적 설명들에 있어서 결정적으로 중요한 경우가 많다. 만약 어떤 사람이 아틀란타가 미시시피주에 있다고 믿으며 또한 미시시피주에 결코 가지 않기를 원한다면 우리는 그 사람이 왜 아틀란타에 결코 가지 않기를 원하는지를 설명할 수 있다. 그 사람은 우리가 형식 논리학의 체계에서 표현할 수 있는 추론을 한 것이다. 만약 우리가 이러한 두 심적 상태들(믿음과 욕구)의 저변에 있는 신경 상태들만을 이용하여 설명을 시도한다면, 우리의 심리학적 설명에 결정적으로 중요한 이러한 명제들간의 논리적 관계는 상실될 것이다. 그것은 우리가 활용할 수 있는 것은 두 신경생리학적 상태들간의 임의적인 관계뿐일 것이기 때문이다. 우리는 신경 정보만 가지고서는 어떤 사람이 이성적인지 아닌지를 평가할 수 없을 것이다. 우리는 잘못된 정보로부터 적절하게 추론한 사람과 참인 정보로부터 비논리적으로 추론한 사람을 서로 구분할 수 없을 것이다. 따라서 만약 우리가 신경과학 이론만을 가지고 있다면 우리는 합리성을 판단할 수 없을 것이고, 이에 따라, 우리는 설명적 힘을 잃게 될 것이다. 그렇다면 신경과학 이론은 심리학 이론보다 어떤 측면에서 빈약한 것이며, 그래서 포더는 우리가 심리학 이론을 신경과학 이론으로 환원하려 해서는 안 된다고 주장한다.

### 3. 강한 자율성 결론의 원리적 문제점과 실용적 문제점

어떤 철학자들은 적절한 교량 법칙들을 구축하는 데에 따른 심각한 어려움들 때문에 심리학과 같은 상위 수준의 과학들이 그에 상응하는 하위 수준의 과학(예컨대 신경과학)으로 환원가능하지 않을지 모른다는 결론을 도출하였다. 예를 들어서 포더는 상위 수준의 과학

들은 하위 수준의 과학들로부터 강한 의미에서 자율적(autonomous)이므로 심리학과 같은 상위 수준의 과학 연구는 신경과학과 같은 하위 수준의 과학 작업과는 독립적으로 발전해야 한다는 더욱 강한 결론을 끌어내었다. 더욱이 두 학문 영역들은 그들의 논의 영역에서의 대상들을 매우 다른 방식으로 분류하므로 하위 수준의 과학들과 상위 수준의 과학들은 서로에 대해 유용한 지침들을 제공할 수 없으며 오로지 그들 각자의 방법 속에서 자체의 문제들을 추구해야만 한다.

그러나 강한 자율성(strong autonomy) 결론은 원리적인 근거와 실용적인 근거 양쪽에서 의문의 여지가 있어 보인다. 원리적 근거 중 하나는 이론 환원 모형이 비판자들의 주장과는 달리 상위 수준의 용어들과 하위 수준의 용어들 사이에 1 대 1 대응 관계를 요구하지 않는다는 것이다. 리처드슨(1979)은 이론 환원 모형의 주된 옹호자인 네이글조차도, 하위 수준의 상태들이 상위 수준의 상태를 왜 실현하는지를 설명할 수 있는 한에서, 높은 수준의 속성이 다중적으로도 실현될 수 있음을 허용하였다고 주장한다.

패트리샤 처칠랜드(Patricia Churchland)는 온도라는 상위 수준의 속성이 항상 동일한 방식으로 실현되는 것은 아니라는 것을 보이기 위하여, 더 기초적인 통계역학의 법칙들로 환원된다고 인정되는 열역학의 사례에 호소한다. 구체적으로 말해, 온도는 기체 상태에 있어서만 분자의 운동 에너지이지, 플라즈마나 고체 상태에서는 그렇지 않다는 것이다. 환원은 각 환원을 논의영역 상대적으로(domain relative) 취급함으로써 수행된다. 처칠랜드는 심리학의 법칙들을 논의영역 상대적으로 신경과학의 법칙으로 환원하는 것이 위 온도의 예와 마찬가지로 받아들일 만하다고 주장한다. “만약 인간의 두뇌와 전자적 두뇌 양자가 특정한 유형의 인지적 조직(organization)이라면, 우리는 분명히 구분되는 두 개의 논의영역 상대적인 환원들을 갖게 될 것이다.” (처칠랜드, 1986, 357쪽)

처칠랜드는 또 심리학의 범주들이 신경과학의 범주들과 상호 상반된 목적을 가진 것처럼 보인다는 사실은 신경과학에서의 이론 발전



의 현행 수준이 반영된 것일 뿐이라고 주장한다. 현재까지는 신경과학에서의 이론작업은 매우 낮은 수준에, 예를 들어서 단일 뉴론의 수준에 초점을 맞춰왔다. 하지만 신경과학은 점점 더 유기체의 보다 높은 수준에 초점을 맞추어가고 있으며, 뉴론들의 모임체가 수행하는 정보 처리과정이나 저장과 같은 것들을 관찰하고 있다. 여기서 처칠랜드는 신경과학이 심리학의 목적과 양립 가능한 개념적 장치를 개발할 것이라고 기대하는 것은 전적으로 합리적이라는 입장을 견지한다. 이 주장을 뒷받침하기 위하여 그녀는 신경 체계에 의해 상위 수준의 정보 처리가 수행되는 세 가지의 이론적 작업들들을 기술한다. 소뇌에 의해 수행되는 감각 운동 제어(sensorimotor control)를 설명하기 위하여 개발된 Pellionisz와 Llinas(1982, 1985)의 텐서 네트워크 이론, 주의집중 과정(attentional processes)에 대한 크릭(1984)의 모형, 그리고 지각에 대한 연결주의 또는 병렬분산처리 모형(McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986)이 그것이다.

강한 자율성 입장에 대한 이러한 원리적인 반대에 덧붙여서 실용적인 반대도 있다. 심리학을 전적으로 신경과학으로부터 자율적인 것으로서 해석하는 사람들은 신경과학 연구에서 어떠한 이익도 얻을 수 없을 것이라고 그리고 신경과학 작업을 위한 어떠한 지침도 제공할 수 없을 것이라고 생각한다. 유사한 자율성 주장이 이전에 다른 과학들에서도 제기되었으며 역사적 증거는 그러한 자율성을 고집하는 것이 그 학문분야들에게 해롭다는 것을 보여준다.

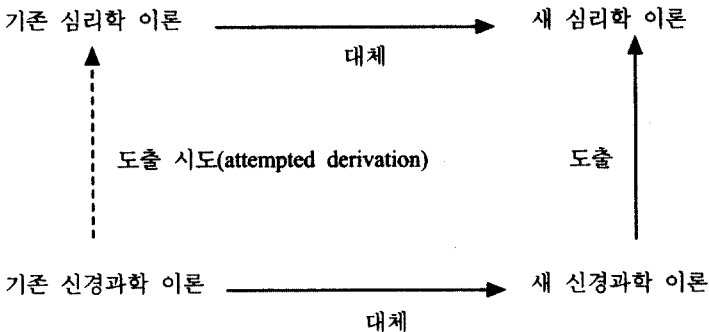
그러한 강한 자율성을 견지함으로써 생겨나는 오류의 예를 리비히(Liebig)의 영양분 처리과정(nutritional process) 모형이나 골(Gall)의 골상학 등에서 찾아볼 수 있다. 리비히나 골(Gall)은 자신들의 초점을 한 학문분야에만 국한시키고, 결과적으로 유용한 지침을 제공할 수 있었던 다른 학문분야에서의 정보를 활용하지 못했던 것이었다. 심리학이 강한 자율성을 고집하는 것은 신경과학으로부터 얻거나 신경과학에 줄 수 있는 유용한 지침과 쓸모 있는 정보를 자신에게서 제거하는 셈인 것이다.

#### 4. 공진화(co-evolution)를 촉진시키는 것으로서의 환원

심리학과 신경과학의 사례에서 이론 환원 모형에 대한 보다 심각한 반대는 인지심리학 이론들을 신경과학 이론들로 환원할 수 있는 실질적 가능성이 현 시점에서는 요원해 보인다는 것이다. 두 학문분야에서 상세하고 완전히 형식화된 이론들이 정립되기 전에는 한 이론이 다른 이론으로부터 도출될 수 없다. 하지만 더 근래에 많은 철학자들이 이론 환원 모형의 또 다른 역할을 제안하였다. 그들 중에서도 워자트(1976a, 1976b, 1979)와 처칠랜드(1986)는, 환원이 단지 이미 개발된 이론들을 과거로 소급하여 결합하는 것에서가 아니라 미래의 이론 발전에서 어떤 역할을 할 수 있다고 제안하였다. 그들의 생각은 만약 이론 환원이 궁극적 목표라면 우리는 이론들 자체를 개발하는 과정에서 상위 수준의 학문분야들과 하위 수준의 학문분야들 양자로부터의 정보를 모두 활용해야만 한다는 것이다. 그리고 만약 이론들이 이런 방식으로 공진화한다면 일단 이론들이 발전될 경우 이론 환원은 자연스레 달성된다는 것이다.

이러한 공진화적인 청사진은, 대부분의 환원 사례들에 있어서 환원되는 이론은 원래의 상위 이론이 아니라 그 상위 이론의 수정된 형태라는 사실에 의해서 뒷받침된다. 예를 들어, 통계역학으로 환원된 것은 원래의 현상적인 열역학 이론이 아니라 그 수정판이었다. 유사하게 분자유전학으로 환원된 것은 멘델(혹은 모건)의 유전이론이 아니라 그 수정판이었다. 이러한 특성들을 반영하기 위하여 샤프너(1967)는 원래 실증주의자들이 제안했던 것보다 더욱 일반화된 형태의 이론 환원 모형을 개발하였다. 이 모형에 따르면 환원은 두 단계의 과정으로 간주된다. 첫 번째 단계는 하위 수준의 이론(과 더불어 교량 법칙과 경계 조건들)으로부터 수정된 상위 수준의 이론을 도출하는 것이다. 두 번째 단계는 (비록 그 원래 이론 중 추후 개선된 부분에 있어서는 실패했지만) 원래의 상위 수준의 이론이 왜 지금까지 그런대로 잘 작동해 왔는지를 우리가 충분히 이해할 수 있을 정도로,

수정된 상위 수준의 이론이 원래의 상위 수준의 이론에 근사적으로 접근하고 있음을 보여주는 것이다. 비록 샤프너가 그 가능성을 논하지는 않았지만 새로운 하위 수준의 이론이 개발되는 것은 당연히 가능하며 뿐만 아니라 그런 경우에 우리는 어떻게 새로운 하위 수준의 이론이 그 선행 이론과 비교되는지를 보여줄 필요가 있을 것이다. 이러한 샤프너 모형의 확장판을 심리학과 신경과학의 경우에 적용함에 있어서 최종 목적은 현행의 심리학 이론을 현행의 신경과학 이론으로 환원하는 것이 아니라 미래의 심리학 이론을 미래의 신경과학 이론으로 환원하는 것이다(아래 그림 참조).



샤프너 자신의 설명에 의하면 환원에서의 두 단계들은 최종 이론들이 개발된 후에 수행된다. 환원을 통해서 우리는 상위 수준의 새로운 이론과 하위 수준의 기존 이론이 어떻게 서로 관련되는지를 알 수 있게 된다. (위 그림과 같은) 공진화적인 모형을 창출하기 위해서는 샤프너의 모형을 수정하여 하위 수준의 (잠재적으로 새로운) 이론과 정합적인, 상위 수준의 새로운 이론을 개발하는 과정을 이론 발달 과정의 한 단계로 해석할 필요가 있다. 더욱이 우리는 환원이라는 용어가 두 활동을, 즉 하위 이론에 대한 상위 이론의 관계에 관련된 활동과 이전 이론에 대한 새 이론의 관계에 관련된 활동을, 포괄하고

있음을 알 필요가 있을 것이다. 따라서 우리는 실제로 두 단계의 과정을 갖게 된다. 첫째 단계가 옛날 이론과 그에 대한 새 이론 사이에 이루어지는 것으로, 기존의 상위 수준의 이론을 새로운 상위 수준의 이론으로 대체하는 것이고, 둘째 단계가 새로운 심리학 이론을 신경과학 이론으로 수준간(interlevel) 환원을 하는 것이다. 이러한 단계들을 구분하는 한 이유는 주어진 수준에서 새 이론의 개발을 촉진하는 요인들이 수준간의 관계들을 수용하는 것뿐만 아니라 그 수준 내에서의 적절한 설명을 개발하는 것과도 관련이 있기 때문이다.

이때 생기는 의문점은, 두 수준의 이론들 사이에서 한 이론이 다른 이론의 발전에 어떤 역할을 할 것인가? 하는 점이다. 여기서 워자트(1976a, 1976b, 1979)는 그 이론 발전의 보다 이른 단계에서 하위 수준의 과정과 상위 수준의 과정들(혹은 그러한 과정들을 지칭하는 용어들)을 연결하는 교량 법칙들을 개발하는 것이 중요하다고 주장한다. 이 단계에서의 목적은 상위 수준의 법칙들을 하위 수준의 법칙들로부터 연역하는 것이 아니라 두 이론들에 의해서 제기되는 주장들간의 불일치점들을 확인하기 위한 것이다. 예를 들어, 세기가 바뀔 무렵에 Boveri와 Sutton은 멘델 유전자와 염색체를 연관시킬 수 있는 증거를 확보하였다. 그 이전에는 염색체를 세포 분열에서 어떤 역할을 담당하는 것으로 보이는 핵 속의 염색가능한 구조체로만 간주하였으나, 염색체를 멘델 유전자와 연관시킴으로써 새로운 교량 법칙을 개발한 것이다. 그리고 나서 그들은 염색체로부터 생겨날 수 있는 이론들과는 달리 멘델의 유전자에 대하여 어떤 주장을 할 수 있는가에 대해 초점을 맞추었다. 이러한 차이점들이 주어지자 각 수준에서 제공된 정보는 다른 수준에서의 탐구들을 안내하였다. 이에 따라서 각 수준에서 개발된 이론들에 대한 수정이 이루어졌다. 여기서 교량 법칙을 통한 수준 상호간의 동일시는 가설들의 산출에 기여하였으며 그리하여 이론 발전에 생산적 역할을 담당하였다.

패트리샤 처칠랜드와 같이 심리학을 신경과학으로 환원하는 것을 옹호하는 사람들은 이 두 학문분야의 발전을 진행시키는 유사한 통

합과정을 예견한다. 처칠랜드는 심리학과 신경과학간의 공진화적 연구 프로그램이 성공적일 것이라고 기대하는 여러 가지 이유들을 제시한다. 첫째로, 우리는 심적 과정들이 실제로 우리의 두뇌 속에서 일어나는 과정들이라고 가정하며 심리학자들과 신경과학자들 모두는 학습이나 기억과 같이 마음-두뇌에서 발생하는 밀접하게 연관된 종류의 활동들을 탐구하고 있다는 것이다. 둘째로, 어떤 영역에서 우리들의 생각을 발전시킬 때 모형에 의존하는 것은 꽤 일반적인데, 인지심리학의 연구자들이 모형으로 주로 채택하는 컴퓨터들보다는 두뇌들이 심리적 현상을 연구하기 위해서는 보다 나은 모형을 제공한다고 주장한다. 그녀는 또한 우리 두뇌들이 다른 유기체들에서 발견되는 두뇌들로부터 진화를 통해 얻어진 산물이며 진화는 일반적으로 보수적이어서 이미 사용되어온 성공적인 고안물들을 보존하기 때문에 보다 단순한 두뇌들을 가진 다른 유기체들의 작동은 인간에게서 일어나는 정보 처리 과정을 이해하기 위한 유용한 모형들을 우리에게 제공할 것이라고 주장한다.

## II. 수준간 방법론으로서의 기능적 분석

우리가 두 번째로 검토할 방법론은 계산적 기능주의가 채택하고 있는 하향식 기능적 분석 방법이다. 계산적 기능주의 또는 정보 처리 기능주의는 심리학이 신경과학으로 환원됨을 부정하는 데서 출발하였다. 다시 말하면 인지 현상은 신경 상태로 환원되지 않고 독립적인 실재성을 인정받으며 인과적 영향력을 행사한다. 이러한 환원주의의 비판의 핵심은 인지 현상의 복수 실현 가능성으로, 기능주의자들은 자신들이 심리 상태와 동일시하는 기능적 상태가 무수히 많은 다양한 물리적 체계들에 의해 실현될 수 있다고 주장한다. 퍼트남(Putnam)과 포더(Fodor)의 이 논변에 따르면 정신적 속성은 신경과 같은 어떤 단일한 물리적 속성으로 환원될 수 없는 독자성을 가지며,

심리학은 자율적인 특수과학으로 성립할 수 있는 형이상학적 토대를 지닌다. 따라서 심리학은 신경과학과 같은 하위과학으로 환원되지 않고 독립적인 자율성을 확보하게 된다. 그러나 앞서 우리는 강한 자율성 결론의 문제점들을 살펴보았다. 따라서 여기서는 강한 자율성 결론을 전제하지 않는 것으로서의 하향식 기능적 분석 방법에 대해 살펴보자.<sup>1)</sup>

### 1. 기능적 분석에 대한 목적론적 견해와 기계론적 견해

기능적 분석에 대해서는 최근에 제시된 두 가지 유력한 견해가 있다. 그중 하나는 라이트의 제안을 토대로 하여 기능을 진화의 과정에서 선택된 효과로 보고자 하는 견해이고, 다른 하나는 기능을 어떤 체계가 가지는 역량에 대한 분석적 설명에서 사용되는 하부 역량으로 이해하고자 하는 커밍스의 견해이다. 이 두 견해는 기능적 설명의 대상에 대해서도 의견을 달리 한다. 즉, 라이트의 견해에서는, 기능을 수행하는 특질이나 그러한 특질을 가지는 대상의 존재가 설명의 대상인 반면, 커밍스의 견해의 경우는, 기능에 해당하는 하부역량들을 토대로 성립하는 어떤 체계의 역량이 설명의 대상이다.<sup>2)</sup>

기능에 대한 라이트의 분석에 따르면, “엽록소의 기능은 광합성을 하는 것이다”는 주장이 의미하는 것은 “엽록소는 광합성을 하기 때문에 존재하며, 광합성은 엽록소의 존재로부터 비롯된 것이다”가 된다. 이러한 라이트의 견해는 어떤 특질이나 특질을 가지는 대상(엽록소)의 존재를 설명하기 위해 목적이나 기능(광합성)에 호소한다는 점에서 목적론적 견해이다. 또 기능에 대한 라이트의 견해는 기능을 ‘진화의 과정에서 선택된 효과’로 본다는 점에서 진화론을 배경으로

- 
- 1) 강한 자율성 결론을 받아들일 경우, 기능주의는 이론상 물리주의뿐만 아니라 이원론과도 양립가능하다.
  - 2) 이 두 가지 기능적 분석의 관계에 대한 자세한 논의는 “인지과학의 방법: 기능적 분석”(조인래, 『철학사상』 제12호, 2001년, 189-219쪽, 서울대학교 철학사상연구소) 참조.

하여 나온 것으로 이해할 수 있다. 19세기 중반 다윈은 종의 기원을 설명하는 과정에서 신의 역할을 제거함으로써 결과적으로 과학적 유물론의 위상을 강화하는 데 결정적인 기여를 하였다. 그러나 그가 한 일은 목적이나 기능에 의존하지 않고 종의 기원을 설명한 것이 아니었다. 다윈의 진화론에서 종의 유지나 분화는 개체가 자손들을 번식하는 과정에서 무작위적으로 발생하는 변이와 각 개체의 표현 형질들에 대한 자연 선택의 기제를 통해 일어난다. 결국, 기존의 창조론에서 신이 하던 역할을 다윈의 진화론에서는 자연이 수행하게 되는 중요한 변화가 일어난다. 그러나, 다윈은 이 과정에서 기능과 같은 목적론적 개념을 제거하기보다는 그것이 자연주의적 틀 속에서 어떻게 유지되고 사용될 수 있는가를 보여 주었다.

이와 같이 진화론을 자연주의적 목적론으로 이해하면, 목표 기여적(goal-serving) 장치는 자연 선택 과정의 산물로 볼 수 있다. 즉, 목표 기여적(goal-serving) 장치는 이전 유기체에 그 장치가 있었다는 것이 그 생물체의 생존에 도움이 되었기 때문에 현재의 유기체에 현존한다는 것이다(Wimsatt, 1972; Wright, 1976). 여기서 간과하지 말아야 할 점은 그러한 선택이 그 목표 기여적 장치의 수준보다 높은 수준에서 작동한다는 것이다. 예컨대 간 기능의 선택은 간을 가진 생물체에서 이루어진다. Machamer(1977)는 목적론적 논의를 야기하는 이러한 수준간의 관계가 이론 환원이 연결하는 것과는 아주 다른 방식으로 수준간을 연결시킨다는 것을 인식했다. 목적론적 현상을 설명하는 데 있어서, 한 수준에서 일어나는 일에 대한 설명을 다른 수준에 대한 분석으로부터 도출해 내지는 않는다. 오히려 연구자는 한 수준의 현상이 어떠한 방식으로 더 높은 수준의 처리과정들과 맞물리고, 그에 의해 영향 받는지를 보여준다. 그러므로 목적론적 현상은, 한 수준에서의 어떤 현상이 상위 수준의 맥락에 어떻게 위치하는지, 그리고 그 맥락에 그 현상이 기여하는 것 덕분에 어떻게 현존하는지를 보여주는 수준간 이론에 의해 설명될 수 있다(Bechtel, 1985).

그러나 실제로 인지과학 방법론으로서의 '기능적 분석'과 관련하여

목적론적인 라이트의 견해가 지니는 직접적인 의의는 별로 크지 않다. 그 의의를 찾아본다면, 기능주의자들이 인간 마음의 정보 처리 능력을 진화 과정상의 선택된 효과로 보는 관점을 취할 수 있다는 것이다. 라이트의 분석에 따르면, “인간의 기능은 정보 처리를 하는 것이다”라는 주장이 의미하는 것은 “인간은 정보 처리를 하기 때문에 존재하며, 정보 처리 능력은 인간의 존재로부터 비롯된 것이다”가 된다. 그리고 인간의 정보처리 기능은 인간의 진화과정에서 선택된 효과로 이해되며, 따라서 인간의 정보처리 기능이 인간의 생존에 도움이 되었기 때문에 인간은 정보처리 기능을 가지고 있다는 것이다. 어쨌든 인지과학 방법론으로서의 ‘기능적 분석’을 말하면, 그것은 일반적으로 커밍스류의 기계론적 견해를 지칭하는 것이다.

## 2. 기계론적 설명으로서의 기능적 분석

심리학과 신경과학간의 관계는 커밍스류의 기능적 분석의 관점에서 살펴보면, 이론 환원 모형에 따라 고려할 때와는 상당히 달라 보인다. 우리가 여기서 그 관계를 기능적 분석의 관점에서 살펴보는 것은 수준간 이론의 특별한 유형인 기계론적 설명(mechanistic explanation)을 개선하기 위한 시도로 간주될 수 있다. 그것은 인지가 특정한 유형의 기계에서 어떻게 야기되는가를 보여줌으로써 인지를 이해하려는 긴나 긴 노력에 부합한다. 기계론적 입장을 취하는 연구자들은 마음을 기계에 비유하여 왔는데, 초기에는 증기기관과 전화 교환기가 유비의 사례를 제공해준 반면, 오늘날에는 컴퓨터가 그 역할을 대신하고 있다. 그러나 기계론적 연구의 목표는 여전히 수 세기동안 그대로이다. 즉 알려진 기계들에서 발견되는 처리과정들로부터 인지가 어떻게 야기될 수 있는지를 보임으로써 인지를 이해하려는 것이다.

최근까지 인간이 만든 기계에 적용되어 온 기본적 사실증의 하나는, 기계들은 기계 전체의 과제에 종속적인 하부과제(subtask)들을 담당하는 부분들이 수행하는 바에 따라 전체적인 행동이 이루어진다는



것이다. 우리가 익숙하지 않은 기계에 마주쳐서 그것이 어떻게 작동하는지를 설명하려 할 때, 우리는 일반적으로 그것을 분해하여 그 부분들이 어떻게 작동하는지를 보고, 그리고 난 후 서로 다른 부분들로부터 기계를 재조립하려 할 것이다. 인지 체계와 같은 자연의 체계를 기계처럼 간주하는 연구자들도 똑같은 일을 하려 할 것이다. 그들은 그 인지체계를 분해하여 그들의 구성요소를 결정하고, 이러한 구성요소들이 어떻게 기능하는지를 보고 나서는, 인공적인 부분들로부터 새로운 기계를 조립하거나, 그렇지 않으면 그 부분들이 함께 작동하여 어떻게 이 체계의 행동을 산출하는지를 보여주려 할 것이다. 어떤 한 수준에서 인간 인지 기계의 부분들은 신경구조이기 때문에, 그 기계가 어떻게 작동하는지를 설명하는 과제는 우리를 신경과학으로 인도할 것이다. 그러나 최소한 최초에는 그럴 필요가 없을 수도 있다. 우리는 종종 신경구조의 구성요소가 아니라 그들이 행하는 바의 관점에서 부분들을 기술할 수 있기 때문이다. 종종 심리학자나 언어학자와 같은 인지과학자들을 신경과학자들과 구별하는 것은 단지 그들이 체계의 부분들을 인지적인 기능적 요소들로 보느냐 아니면 신경 구조들로 보느냐이다.<sup>3)</sup> 체계를 규정하는 데 있어서 이러한 서로 다른 방식은 종종 이 구성요소들의 행동을 연구하는 서로 다른 기법의 사용으로 이어진다. 아직은 인지과학과 신경과학 양자를 위한 공통적인 전반적 전략은 그대로 남아 있다. 즉 마음과 뇌의 구성요소들, 그리고 그것들의 상호작용이 관심있는 행동을 어떻게 산출하는지를 보여주는 것이다. 연구자들이 행동을 그것을 산출하는 기제(mechanism)에 연결시키는 한 그들은 수준간 이론들(interlevel theories)을 개발하고 있으며, 이러한 수준들이 서로 다른 분야들(예컨대 인지과학과 신경과학)에서 연구되는 한 그것들을 함께 고려하는 것은 분야들을 연계시키는 이론으로 귀결될 것이다.

인지과학자들이 행하는 기계론적 연구의 노력에는 몇몇 두드러진

---

3) 보통 인지과학을 구성하는 한 분야로 신경과학도 포함되나, 여기서는 인지과학을 신경과학과 대비시켜 좁은 의미로 쓰이고 있다.

특성들이 있다. 우선, 대부분의 기계론적 연구의 기저에는 어떻게 자연이 조직화되어 있는가에 관한 가정이 있다. 기계론적 연구자들은 자연은 Simon(1980)이 지칭한 바 “분해가능한(decomposable)” 방식으로 조직되어 있다고 가정한다. 즉, 체계의 부분들은 거의 자율적인(autonomous) 방식으로 작동하는 모듈(module, 단위체)들로 해석 가능하다는 것이다. 주어진 모듈은 다른 모듈로부터 입력을 받아들이고 이어서 또 다른 모듈에게로 출력을 내보내며, 그 과정 중에 입력에 대한 자체의 조작을 수행한다. 자신의 과제를 수행하는 각 모듈의 능력은 주로 그 내적 속성에 기인한다. 모든 체계가 실제로 거의 분해 가능한 방식으로 설계되어 있다는 가정은 틀리기 쉬운 것이나, 아주 최근까지 우리는 기계를 어떻게 만들어야 하고 자연적인 체계를 어떻게 연구할 것인지에 대해, 이러한 가정을 하지 않은 모형을 가져 보지 못했다.

### 3. 직접적 국소화와 복잡한 국소화

실제의 연구에서, 기계론적 설명을 개발하는 연구자들은 판단을 내려야 하는 여러 중요한 상황들에 직면하게 된다. 첫 번째로 내려야 할 판단은 물음의 대상이 되는 현상에 대한 제어가 이루어지는 장소(locus of control for the phenomenon)가 어떤 체계에 위치하고 있는지를 자연에서 알아내는 것이다. 실제로 이것은 종종 논쟁거리가 되는 판단이다. 인지주의와 행동주의의 대립은 상당부분 마음-뇌가 행동의 제어가 이루어지는 장소로서 취급될 수 있느냐에 대한 대립이고, 또 행동에 대한 제어가 환경으로부터의 강화를 포함하는 더 확장된 체계에 위치해야 하는지에 대한 대립이다. 어떠한 체계가 제어의 장소인지에 관한 결정은, 우리가 최소한 체계의 경계를 확정짓고 나서야 체계를 분해하기 시작할 수 있기 때문에 중요하다. 그러나 인지과학의 연구자들은 일반적으로 마음-뇌를 제어의 장소로서 결정하고 있다.

주어진 체계가 어떻게 제어의 장소가 되는지를 설명하기 위해서는

체계를 분해하고 그 체계의 부분들이 행동을 어떻게 야기하는지를 보여주는 것이 필요하다. 일반적인 첫 번째 전략은, 그 자체로 과제를 수행하는 단일한 구성요소를 체계 내에서 찾아내는 것이다. Bechtel과 Richardson이 “직접적 국소화(direct localization)”라고 지칭한 이러한 전략은 Gall과 Broca와 같은 주목할 만한 연구자들의 탐구에서 그 사례를 찾아볼 수 있다. Gall은 사람들 사이에 나타나는 성격상 특징의 차이점들을 단위능력들을 가정함으로써 설명하려 했는데, 그 단위능력의 기능 각각이 하나의 성격상의 특징을 담당한다는 것이다. 성격상의 차이점들은 관련된 단위능력의 발달에 있어서의 차이점에 기인한다. 또한, Gall은 단위능력과 관계된 특징들의 표현을 두개골상의 돌출부와 연관지음으로써 이러한 단위능력들을 뇌의 영역과 동일시하려고 했다. 이리하여 Gall은 체계의 활동을 그 체계의 부분들의 특성과 연계시키는 직접적 국소화 연구의 주된 전략의 사례를 보여주고 있다. Gall의 연구의 기저에는 행동적 특성이 체계의 구성요소에 기인한다는 강한 단위성(modularity)의 가정이 있다. 일단 이러한 가정이 이루어지면 연구전략은 아주 자연스러워진다. 연구자는 어떤 구성요소가 특징들을 담당하는지를 보여줄 관련된 증거를 찾아야 한다. 이와 같은 증거를 찾아내는 것이 기계론적 연구 프로그램의 특징을 결정하는 두 번째 주된 판단을 구성한다.

Gall이 종종 사기꾼이라 비난받음에도 불구하고, 체계의 활동들을 그 체계의 구성요소들의 특징과 연계시키려 시도하는 전략 자체는 다른 사람들에게도 널리 받아들여져 왔다. 예를 들어, Broca(1861)는 대뇌 피질의 왼쪽 측두엽 영역이 언어 생성 기능을 담당한다는 그의 주장을 전개하는 데 있어서, 비록 다른 종류의 것이긴 하지만 Gall과 마찬가지로, 영역에 대한 손상과 기능의 결손간의 연계적 증거에 의존하였다. 종종 이러한 연계적 기법은 손상(lesion)을 관찰하여 기능의 결손을 보여주는 실험적인 절차로 확장되었다. 손상 기법에 대한 대안으로는 자극 절차(stimulatory procedures)가 있는데, 이것은 특정한 기능이 위치하고 있다고 여겨지는 영역을 자극하여 그것이 문제

의 그 기능의 수행에 결과를 미치는지 보는 것이다.

손상이나 자극 절차는 손상과 기능결손의 연계 또는 자극과 수행의 연계와 같은 단순한 연계를 입증하는 데 주로 의존하고 있다. 그러한 기법이 종종 연구에서 등장하기는 하지만, 두 기법 모두 신뢰성이 높지는 않다. 손상과 기능결손의 연계는 문제의 장소가 그 자체로 기능의 수행을 담당하는지 아니면 단지 부수적인 역할만을 할 뿐인지를 입증하지 못한다. 유사하게, 자극과 향상된 수행의 연계 또한 그 자체로는 그 장소가 해당 기능이 수행되는 곳이라는 것을 입증하지는 못한다. 왜냐하면 그 영역에 자극을 가하는 것이 단지 다른 곳의 활동을 초기화하는 데 기여하는지도 모르기 때문이다. 그러나 아마도 더욱 중요한 것은 직접적 국소화는 아무리 정확하다 해도 그 자체가 기계론적 설명을 제공해주지는 않는다는 것이다. 기계론적 설명은 한 체계가 어떻게 기능을 산출할 수 있는지 보여줄 것을 요구하나, 직접적 국소화는 그것을 시도조차 하지 않는다. 그것은 단지 기능을 담당하는 하위 체계를 식별해 내려할 뿐이다.

기능에 대한 기계론적인 설명은 최소한, Bechtel과 Richardson이 기능의 “복잡한 국소화(complex localization)”라고 지칭한 것을 요구한다. 이것은 체계가 분해 가능하다는 가정뿐 아니라 서로 다른 구성요소들이 서로 다른 하위 과제들을 수행한다는 가정을 요구한다. 이에 따라 ‘복잡한 국소화’ 설명은 두 가지 작업을 요구한다. 그 하나는 전체 과제를 구성요소 과제로 기능적으로 분해하는 것이고, 다른 하나는 이러한 과제들 각각을 수행하는 구성요소들이 있음을 예시하는 것이다. 이러한 분해법을 개발하기 위해 연구가들은 전체 과제가 어떻게 여러 구성요소들에 의해 각각 수행되는 단순 과제들의 집합적 수행으로부터 야기될 수 있는지를 해결해야만 한다. 심리철학에서 인지적 활동을 구성요소의 활동들로 분석하는 것을 지칭해 종종 난쟁이 분석(homuncular analyses)이라고 하는데, 그것은 마음의 내부에 작은 주체들(난쟁이들)의 무리가 있어서 각자 작은 과제를 수행한다고 가정하기 때문이다. 그러한 분석은 어떻게 사람들이 지능적으로

행동하는가를 설명함에 있어서 또다시 지능적인 주체들을 가정하지 않느냐는 물음의 여지가 있어 보이지만, 그러한 순환성은 가상의 난쟁이는 전체 체계보다는 덜 지능적이고 더 단순한 과제를 수행한다는 것을 요구함으로써 피해갈 수 있다(Bechtel, 1988a; Dennett, 1978; Lycan, 1981). 두 번째 요구조건은 난쟁이 분석에서 상술된 과제들을 수행하는 체계내의 구성요소들이 있다는 것을 경험적으로 예시하는 것이다. 여기서 연구자들은 직접적 국소화에서 등장한 바와 같은 기법들, 즉 구성요소의 행동을 체계의 행동상의 변화와 연계시키는 기법들을 사용해야 한다. 마찬가지로 여기서도 대부분의 두드러진 기법들은 손상과 자극 절차이다. 이 경우에 있어서의 차이는 전체 행동상의 변화가 단순히 체계의 전체 행동이 차단되거나 촉진되는 형태로 나타나는 것이 아니라 그 전체 행동의 다양한 변형태로 나타난다는 것이다. 이것은 기능적 분석으로부터 이끌어 낸, 행동이 어떻게 변형될 것인가에 관한 예측에 상응하는 것이다.

역사적으로 그러한 복잡한 국소화 설명이 개발되어 온 다양한 방식이 있으나, 그중 한 모범적 사례를 현재 진행되고 있는 인지신경과학 작업에서 찾아볼 수 있다. O'Keefe와 Nadel(1978)은 두 종류의 공간적 기억 체계에 대한 증거를 밝혔는데, 그 중 하나는 객관적인 환경지도를 포함하는 것인 반면, 다른 하나는 한 장소에서 다른 장소로 이동하는 방향들의 집합에 의존하는 것이다. 그들은 또한 손상 연구로부터 객관적 지도를 처리하는 능력이 해마 영역에 국소화되어 있다는 증거를 찾아냈다. 이것은 맵핑 기능을 사용한 직접적 국소화 작업에 해당한다. 그러나 O'Keefe와 Nadel은 해마가 어떻게 이러한 기능을 수행하는지를 설명할 필요를 인식했다. 이 요구를 충족시키기 위해, 그들은 해마를 형태학적, 신경생리학적인 바탕 위에서 서로 다른 부분들로 분해하였다. 그리고 나서 그들은 해마의 서로 다른 부분들에서의 활동이 어떻게 해마에 국소화된 전체 과제인 정신지도 처리능력을 설명해 줄 수 있는지에 대한 분석을 전개했다. 이 경우의 의미있는 특징은 O'Keefe와 Nadel이 기능의 직접적 국소화로부터 출

발하여 이 기능이 어떻게 수행되는지 설명하기 위해 낮은 수준으로 내려갈 것을 결심했다는 것이다. 그들은 해마 영역에서의 신경세포의 활동에 관한 정보를 이용하여 해마가 이러한 기능을 어떻게 수행하는지에 대한 난쟁이 분석을 전개해 나간 것이다.

#### 4. 기능적 분석과 환원

이러한 짝막한 묘사가 암시하는 바와 같이, 기계론적 설명을 개선하는 방식을 결정하는, 연구자가 내려야 할 다양한 중요한 판단들이 있다. 여기서 우리가 주목할 필요가 있는 것은 그러한 판단들의 산물이 가지는 특성이다. 그 산물은 이론 도출이 아니고 기능적 분석을 통한 분야 연계 이론이다. 연구자들이 산출하는 것은 어떤 기능이 체계내의 부분들이라는 관점에서 어떻게 수행되는가를 분석하는 것이다. 연구자들은 한 이론을 다른 이론으로부터 도출해 내려고 하지 않으며, 낮은 수준에서 특정한 종류의 작동이 어떻게 하여 높은 수준의 현상을 야기하는가를 예시하려고 한다. 높은 수준에서 연구된 현상들(예컨대 공간 기억)이 낮은 수준의 활동의 산물이라는 이유로 상위 수준의 과학이 하위 수준의 과학에 종속되지 않는다. 기능적 분석을 통한 분야 연계 이론은 그것들이 어떻게 산출되는가를 보여주며, 따라서 이론화의 두 수준을 연결한다.

더욱이 그러한 분야 연계 이론들을 수준간에 걸쳐 개발하는 과정은 각 수준에서의 이론을 개발하는 데 있어 중요한 역할을 하는 경향이 있다. 예컨대, 두 유형의 공간 기억 체계간의 구분은 해마 영역에서 어떤 종류의 기억을 찾아야 하는지에 관해 알려줌으로써 O'Keefe와 Nadel의 작업을 안내했다. 해마의 구조에 대한 정보는 이번에는 공간 기억의 과제들을 수행하는 데 관련되는 처리과정들의 유형에 관해 안내했으며 연구자들로 하여금 처리과정에 대한 새로운 난쟁이 분석으로 이끄는 도구가 되었다.

이전 장에서, 필자는 심리학을 신경과학으로 이론 환원 모형에 의

거하여 환원하는 데 대해 반대하는 사람들이, 인지적 이론들의 신경과학 이론으로부터의 전적인 자율성이야말로 환원에 대한 유일한 대안이라고 어떠한 방식으로 주장해 왔는지 기술하였다. 여기서 묘사한 기계론적 기능적 분석 모형은 다른 가능성을 열어준다. 이 모형은 인지적 연구와 신경과학적 연구간의 상호작용을 환원 없이 수용할 수 있다. 그것은 양쪽의 분석으로 하여금 분야 연계 이론을 개발하기 위한 시도 안에서 서로 정보를 주고받을 수 있게 하지만, 한 설명을 다른 설명 하에 종속시킬 것을 요구하지는 않는다. Von Eckardt(1978)는 유용한 유비를 제공하였다. 우리는 특정한 오믈렛 제조기계가 오믈렛을 어떻게 만드는지에 대해 기능적 분석을 개선할 수 있다. 그러나 우리의 분석에서 단계들은 오믈렛 기계가 실제로 수행하는 단계와 상응할 필요는 없다. 우리가 잘못된 기능적 분석을 하고 있다는 것을 발견하는 하나의 방법은, 우리의 기능적 분석에서 확인된 다양한 과제들을 수행하는 구성요소가 체계 내에 없다는 것을 발견하는 것이 될 것이다. 그러나 그것이 단지 잘못된 기능적 분석을 철회하게 하는 데에만 기여하는 것이 아니다. 기능을 수행하는 구성요소들을 확인하려는 시도는 기능적 분석을 개발하는 방법에 관한 심화된 통찰을 제공할 것이다. 부분들이 어떤 일을 하는가로부터 시작하여, 우리는 체계가 실제로 어떻게 작동하는가에 관한 새로운 개념화에도 인도될 것이다.

이론 환원 모형과는 달리, 기능적 분석 모형은 낮은 수준이 최종 평결을 내리길 요구하지 않는다. Von Eckardt가 제공한 또 다른 예는 어떻게 높은 수준이 낮은 수준의 고려에 우선할 수 있는지를 보여준다. 그녀는 보통 실인증 환자(agnosics)라고 분류된 사람들 중 특정한 부류의 환자들은 단지 이름 붙이는 작업의 결손만을 가지고 있을 뿐인 것으로 재분류되어야 한다고 주장하며, Geschwind와 Kaplan(1962)의 논변을 비판하였다. Geschwind와 Kaplan은 대상 인식의 중심부를 가정하는 것은 불필요하며, 따라서 단순화의 이유로 두 결손을 하나로 재분류해야 한다고 주장하였다. Von Eckardt는 데이터가 대상 인

식 중심부를 가정하는 더 고전적인 모형에 반하지 않는다는 것과 더불어, 대상 인식의 과정이 심리학의 여타 정보처리적 설명에서 중심적으로 등장하기 때문에 대상 인식 중심부도 받아들여져야 한다고 주장하였다. 그러한 중심부를 가정함으로써 보통의 사람들이 행동하는 방식에 대한 더 나은 설명을 제공할 수 있기 때문이다. 그러므로 체계의 전체적인 기능을 설명하는 기능적 분석의 능력이 특정한 손상의 결과에 대한 해석을 평가하는 데 고려되고 있는 것이다.

기능적 고려가 오직 신경과학적 탐구에만 기반하여 만들어진 주장을 재평가하도록 초래한 다른 유익한 경우는, 언어 처리에 대한 브로카와 베르니케의 영역이 갖는 기여를 최근에 재해석한 데서도 발견할 수 있다. 전통적인 해석은 베르니케 영역은 언어 이해의 중심부이며 브로카 영역은 언어 생성을 담당하는 것이라고 주장했다. 이 해석은 브로카 영역의 손상은 언어 생성에서 가장 명확히 나타나는 반면 베르니케 영역의 손상은 이해를 무력하게 한다는 사실에 기초했을 뿐 아니라, 입력은 그와 관련된 출력 처리가 시작될 때까지는 한 채널에서 처리된다는 관념연합설(associationistic) 모형을 수용함에 근거하였다. 그러나 현대 언어학에서는 분석의 범주로 실어증 연구가들의 이해-생성간 구분에 강조를 두는 것이 아니라 대신 음운학적, 구문론적, 의미론적 그리고 화용론적 분석의 범주를 채용하고 있다. 그러므로 현대 언어학에 훈련된 연구자들은 언어학 이론의 작업틀 내에서 실어증 데이터를 재해석하려 했다.

연구자들은 이러한 재분석을 수행하면서 전통적인 해석과는 정반대로, 브로카의 실어증 환자들이 이해 결손 현상을 나타낸다는 증거를 발견하였다. 이러한 현상은 독특한 통사론적 또는 범주연결어적인(syncategorimatic) 용어들이 문장의 의미에 중요할 때 그리고 이러한 의미가 다른 용어에 의해 중복적으로 나타나지 않을 때 발생한다. 예컨대, “접시 위에 냅킨을 놓아라(Place the napkin on the plate)”라는 문장에서 “위에(on)”라는 단어는 중요한데, 사람으로 하여금 냅킨을 접시 아래에 놓도록 하는 것도 똑같이 가능할 것이기 때문이다. 그러



한 문장들에서 브로카의 실어증 환자들은 이해 결손 현상을 나타낸다. 계속된 조사에서 이 연구자들은 정상적인 피험자들이 구문론상 지시하는 단어(예컨대 “on”)를 보통의 내용을 가진 단어들(예컨대 “table”)보다 훨씬 빨리 처리하는 반면, 브로카의 실어증 환자들은 그것을 하지 못한다는 것을 보여줄 수 있었다(Bradley, Garrett, & Zurif, 1980). 그러므로 이 연구자들은 브로카 영역은 언어 생성이 아니라 구문론적 분석을 담당한다고 제안하였다. 여기서 원래 신경과학적 기초를 둔 실어증 연구로부터 개진된 설명을 수정하는데 언어심리학적 관점이 중심으로 등장하였다. 그러나 더욱 최근에 확장된 연구는 언어학적 관점 자체에 가능한 수정을 지적하고 있는 것으로 보인다. 브로카 영역은 언어적 기능에 특정하게 관련되어 있는 것이 아니라 자동적 기능 일반에 보다 일반적으로 관련되어 있을지 모른다는 제안이 발전되어 왔다(Grodzinsky, Swinney, & Zurif, 1985). 그러한 제안은 실어증 연구의 전통과 Chomsky의 전통 양자의 기반이었던 단위적(modular) 언어 처리 단위라는 생각 자체를 위협하며, 다른 인지적 능력에 언어적 능력을 통합할 것을 제안하고 있다.

여기서 기술한 복잡한 국소화를 통해 기계론적 설명을 개진하는 과정은, 과학의 수준간의 관계를 설명하는 데 있어서 이론 환원 모형에서 제공하는 것에 대해 하나의 대안적인 설명을 제공함을 알 수 있다. 그러나 그것은 다음과 같은 가정 위에 세워진다는 것을 강조하는 것이 중요하다. 자연은 분해 가능한, 단위적(modular) 방식으로 설계되어 있다. 기계론의 프로그램이 지난 세 세기에 걸쳐 생명과학에서 주된 진보를 이루어 왔지만, 그것 역시 종종 비판받아왔다. 생물학에서 기계론에 대한 비판자들은 종종 생기론자들(vitalists)이라 불려 왔으며 반면 심리학적 반대자들은 이원론자들(dualists)이라는 표지가 붙어 왔다. 이런 표지는 사실상 광범위한 서로 다른 입장들을 포함하고 있으나, 그들은 모두 공통적 문제에 직면하였다. 분해의 가정 없이는 어떠한 설명적 연구 프로그램도 가능해 보이지 않는다. 과학자는 어떠한 체계의 행동을 분해하지 않고 기술할 수 있을지는 몰라

도, 왜 그것이 그렇게 행동하는가를 설명하진 못한다. 유일한 가능한 설명은 행동을 특정한 행동에서 자신을 드러내는 생기(vital power)나 비물질적인 마음에 의존하는 길뿐일 것이다.

### Ⅲ. 수준간 방법론으로서의 병렬분산처리

최근에는 전체적인 기능적 작동을 구성요소들의 작동으로 분해할 것을 요구하지 않는 대안적 관점이 대두되었다. 전통적인 인지적 모형은 인지적 과제를 표상을 조작하기 위한 규칙들로 분해해 왔다. 이것은 최근 제안된 병렬분산처리 체계(PDP 체계, 또는 연결주의 체계)에는 적용되지 않는다. 이러한 체계는 매우 단순한 단위들로부터 만들어진다. 이 단위들은 서로에 연결이 되어 서로를 활성화하거나 억제할 수 있다. 이 체계는 열역학적 원리들에 의해 작동하여, 초기 입력이 들어오면 단위들은 서로간의 활성화나 억제를 주고받아 최소의 에너지가 실현된 안정된 상태에 도달한다. 체계에 입력으로서 공급된 활성화 패턴은 체계에 과제를 할당하는 것을 나타내며, 반면 결과로 만들어진 안정된 패턴은 문제에 대한 체계의 해답을 나타낸다. 이 체계의 행동은 전적으로 국지적 작용, 특히 단위들간의 연결에 의해 결정되므로, 전체 과제를 구성요소의 과제로 쪼개는 규칙들에 의해 지배되지 않는다. 전반적 처리과정은 다소간 특정한 인지적 규칙에 상응할지 몰라도, 이러한 규칙들은 체계의 작동에 등장하지는 않는다. 규칙과 유사한 행동은 단지 창발적인 산물일 뿐이다(McClelland & Rumelhart, 1986; Rumelhart, 1984; Rumelhart & McClelland, 1986).

병렬분산처리 체계에서는, 전체로서의 작업이 구성요소들의 작업 수행 결과가 아니므로, 연구자는 설명의 목적으로 체계를 구성요소의 인지적 작용으로 분해할 수가 없다. 체계내의 작동들은 상당히 다른 종류의 것으로 노드간에 흥분과 억제를 전달하는 것이다. 행동은 구성요소에 의해 수행되는 특정한 작용으로부터 야기되는 것이 아니라

많은 수의 매우 단순한 구성요소들의 상호작용으로 만들어진 전체 패턴의 특성으로부터 야기된다. 연구자들이 할 수 있는 것은 그 행동이 어떻게 발생하는가를 단계 단계별로 보여주는 것이 아니라, 결과로 얻어진 체계가 인지적 용어로 규정될 수 있을 정도로 적절하게 행동하는 것을 보여주는 것이다.

그러므로 병렬분산처리 체계는, 보다 전통적인 기계론적 체계가 기초하고 있는 분해의 원리를 침범하며, 보다 전통적인 기계론적 설명에 대한 대안을 제공한다. 만일 우리의 인지 체계가 병렬분산처리 유형의 체계라면, 위에서 개괄한 기계론적 연구전략은 작동하지 않을 것이다. 왜냐하면 기계론적 연구전략은 우리로 하여금 상호작용의 패턴에 초점을 맞추지 않고 불연속적인 구성요소들을 찾아내도록 이끌 것이기 때문이다. 그러나 이러한 실패가 수준간의 연결을 끊어내는 것이 과도되었다는 것을 보여주는 것은 아니며, 단지 다른 유형의 수준간 이론이 요구된다는 것을 보여줄 뿐이다. 수준간 이론은 병렬분산처리 체계가 인지적 용어들로 기술될 수 있는 행동들을 어떻게 산출할 수 있는지를 보여줄 것이다. 이러한 이론들을 개발하기 위한 전략은 전통적인 것과는 다른 것인데, 그 이유는 이해와 같은 인지적 능력이 체계를 분해하는 것으로부터 나오지 않기 때문이다. 오히려 이해는 어떤 종류의 체계가 어떤 방식으로 행동할지를 보여주는 수학적 모형에 대한 이해로부터 나올 것이다. 그러나 그 결과는 여전히 특정 종류의 처리과정들이 어떻게 다른 처리과정들을 야기하는지를 보여주는 수준간 이론일 것이다(Bechtel, 1986).

#### IV. 수준내 방법론으로서의 경계 허물기와 경계에 다리놓기

지금까지 인지과학 방법론으로 살펴본 환원, 기능적 분석, 병렬분산처리 등은 모두 수준간을 연계시키는 방법론으로, 상위수준에 위치

한 인지과학과 하위수준에 위치한 신경과학을 연결시키는 방법론이다. 그러나 학제적인 인지과학 연구 속에서는 심리학, 인공지능, 언어학, 인류학, 그리고 철학 등과 같이 동일하게 상위 수준에 속하는 인지과학내의 학문 분야들간을 연결시키는 방법론 또한 필요하다. 이러한 학문분야는 대략 동일한 수준에서 자연 속에 있는 유기체의 현상들을 다루므로 부분-전체 관계가 얻어지기 어렵다. 따라서 이들 학문분야 중 하나의 이론으로부터 다른 분야의 이론을 도출할 것을 제안하는 것은 설득력이 없다. 얻을 수 있는 유일한 환원 관계는 이들 이론들 중 몇몇이 모두 동일한 낮은 수준의 이론 하에 포섭되는 경우일 것이다. 그러나 인지과학의 창시자들이 그런 종류의 환원 없이도 참여하는 학문분야들간의 흥미 있는 이론적 관계를 개발하려고 계획했다는 것은 명백하다. 이제 동일 수준에 속하는 학문분야들의 이론들을 연계시키는 방법론에 대해 살펴보자.

Abrahamsen(1987)은 학문간의 접촉에 있어 '경계 허물기(boundary-breaking)'와 '경계에 다리놓기(boundary-bridging)'간의 유용한 구별을 소개했는데, 각각은 다른 종류의 이론적 노력을 수반한다. 경계를 허무는 접촉은 한 학문분야의 종사자들이 자신의 학문분야를 해석하는 새로운 길을 찾기 위해 다른 학문분야에 주의를 기울일 때, 또는 한 학문분야의 종사자들이 다른 학문분야 영역의 일부를 자신의 영역으로 가져오려 시도할 때 일어난다. 한편 경계에 다리를 놓는 접촉은, 한 학문분야의 종사자들이 자신의 학문분야 안에서 정의된 문제를 풀기 위해 관련된 학문분야의 작업에서 도움을 받으려 할 때 발생한다. Abrahamsen은 심리학과 언어학간의 상호작용의 역사에서 양 종류의 노력이 모두 발생했고, 중요한 역할을 했다는 것을 보여주려 했다.

### 1. 경계 허물기

대다수의 주의를 끄는 것은 보통 경계 허물기의 사례이다. Blumenthal(1987)은 심리학과 언어학간의 경계 허물기 관계를 보여주

는 커다란 사례 둘을 기술한다. 양자 모두 학문분야의 종사자들이 자신의 이론적 원천이 빈곤해졌음을 발견하고 다른 학문으로부터 이론들을 자기 학문분야로 끌어들이며 그것을 확장하려 했을 때 발생했다. 19세기 후반 자신들의 이론적 원천의 빈곤함에 불만이었던 언어학자들은 Hobart나 Wundt와 같은 연구자에 의해 발전하고 있는 새로운 심리학적 전망에 주의를 기울였다. 여기서 언어학자들은 언어 연구에 대한 경험적이고 양화 가능한 접근법을 개발할 수 있는 희망을 보았다. 그들의 노력 중 일부는 심리학에서 정신적 처리과정을 탐구하는 도구가 어떻게 언어의 구조를 설명하기 위해 도입될 수 있는가를 보여주는 것이었다. 그러나 이러한 학문간 경계 허물기 노력은 부분적으로는 인지적 과정에 관한 심리학 내에서의 다양한 견해들의 결과로, 부분적으로는 행동주의의 번성의 결과로 금세기 초에는 중국에 이르렀다. 그렇지만 이러한 학문간 경계 허물기는 심리학과 접촉이 끝난 지 오랜 후까지 구조주의적 언어학의 형태로 언어학적 연구에 유산을 남겼다.

1960년대 언어학자들은 심리학에 새로운 전망을 불어넣음으로써 심리학으로부터 받았던 혜택을 되돌려주었다. Chomsky는 언어학의 과제들을 재개념화하는 과정에 있었고, 특히 자신의 생성문법 모형을 가지고 언어의 구조주의적 분석에 대한 대안을 제공하였다. 이 시기에 George Miller, Thomas Bever 그리고 David McNeill과 같은 수많은 심리학자들은 행동주의적 심리학의 이론들이 빈곤함을 발견하고 Chomsky의 언어학에서 잠재적으로 심리학에 유용한 작업들을 보았다. Chomsky 문법에서 제안된 변형은 마음에서 발생할 수 있는 작동을 제시했고, 더욱이 그러한 변형이 일어난다고 확증할 데이터를 찾는 것이 가능해 보였다.

현대 인지 심리학은 시초에서부터 Chomsky로부터 빛을 졌을 뿐 아니라 Chomsky도 자신의 작업을 심리학에 대한 기여로 기꺼이 해석했다(Chomsky, 1968). 그의 언어학적 분석에 기초하여 Chomsky는 문법적 지식은 본유적인 것이며 학습되는 것이 아니라는 그의 생득

설(nativistic) 결론과 같은 심리학적 결론들을 끌어낼 수 있었다. 이러한 주장은 부분적으로는 언어학적 증거에 의해 뒷받침되었고(보편 문법의 증거), 또한 심리학적 근거(언어 학습자가 받는 자극은 학습이 가능하기에는 너무 빈곤하다는 논변) 위에서 구축되었다. Chomsky가 심리학에 하려 했던 또 다른 기여는 능력(competence)과 수행(performance)간의 구별이었는데, 그는 그것을 두 학문간에 심화된 연구를 해야 할 책임이 있는 과제로 보았다. 보편 문법은 자연언어의 문법 원리들을 제한하고, 이러한 원리들을 통달하는 것이 언어 능력을 구성한다. 보편 문법은 보편적이면서도 본유적이기 때문에 언어학자는 그 자신의 자원에 의거하여 능력의 이론을 발전시킬 수 있었고, 기억이나 처리 용량으로부터 비롯되는 실제 언어 수행에서 가해지는 제한에 대한 연구는 심리학적 연구자들에게 남겨두었다.

이러한 두 경계 허물기의 사례에서 볼 수 있었던 것은 한 학문분야의 이론을 토대로 다른 학문분야의 이론이 재구성되는 모습이였다. 두 학문분야는 분야들을 연계시키는 이론을 생성하는 데 있어서 서로 상호영향을 주고받는 상대자였다기보다는, 한 학문분야에서 이미 개발된 이론이 다른 분야의 물음에 대답하도록 단순히 확장되었다. 언어학적 이론이 이제 애초에 설계된 것과는 다른 종류의 문제들에 대답하도록 요구되므로(예를 들면 언어 체계의 구조보다는 정신적 처리과정에 관한 물음), 심화된 연구가 요구되는 것은 놀라운 일이 아니다. 그러나 이런 방식으로 학문분야들을 통합하려는 초기의 시도는 결과적으로 실패하였다. Chomsky에 의해 제안된 학문간 프로그램에 대한 관심은 심리학과 언어학 양쪽에서 모두 감소해 갔다. Reber(1987)와 McCauley(1987)가 보여주려 했던 것처럼, 이것은 서로 다른 요구가 언어학과 심리학의 이론들에 부과되었으며, 각 학문분야의 종사자들은 분야들을 연계시키는 이론을 형성하는 데보다는 주로 자신의 학문분야의 요구에 응하는 데 자신들의 작업을 집중시켰다는 사실에 기인한다.

## 2. 경계에 다리놓기

심리학과 언어학간의 교환이 경계에 다리놓기 노력의 일환으로부터 출현한 그림은 다소 다르게 보인다(Abrahamsen, 1987). 여기에는 한 학문 분야의 이론이 다른 학문분야의 이론을 재구성하도록 유도하려는 시도 대신에 오히려 각 학문 분야의 탐구들간에 유용한 연결을 끌어내려는 시도가 있었다. 분야간 노력으로서 심리학과 언어학의 상호작용을 이해하기 위해 우리는 심리학과 언어학 양자가 모두 언어의 면모들을 연구하지만 양자는 언어를 다르게 해석한다는 것을 인식하는 것이 필요하다. 많은 언어학자들에 있어서(Chomsky를 제외하고) 언어는 주로 추상적으로 분석되어야 할 흥미로운 구조를 가진 문화적 산물이다. 심리학자들에게 언어는 유기체에 의해 이해되고 생성되는 어떠한 것이며, 따라서 관심은 그러한 실시간적 처리과정을 연구하는 데 집중되어 있다.

언어에 대한 다른 설명은 이러한 두 목적을 촉진할지도 모르나 이러한 설명들은 각자가 다른 목적을 추구하는 사람들에게 정보를 제공하는 것이 될 수도 있다. 유기체가 언어를 어떻게 표상하고 조작하는지에 관한 정보처리 모형을 개발하기 위해, 심리학자들은 언어에 대한 기술을 필요로 한다. 언어학자의 기술은 유용한 출발점을 제공한다. 그러나 그것은 정보처리적 설명을 제공하기 위해서는 재형식화될 필요가 있을 것이다. Clark와 Clark(1977; 한 사람은 심리학자이고 다른 사람은 언어학자이다)의 연구는 어떻게 심리학적 연구가 언어학에 의해 그려진 언어에 대한 기술로부터 출발할 수 있는지, 그러나 이러한 기술이 심리학적 탐구의 과정에서 어떻게 재형식화될 수 있는지를 예시한다. 그들은 부정(negation)이나 비교급과 같은 현상에 대한 언어학적 설명을, 정보처리 모형에 통합될 수 있도록 처리과정의 형식에 알맞게 재형식화하였다. 그들은 그리고 나서 오류 패턴이나 반응 잠재력과 같은 심리학적 데이터를 사용하여 그 모형의 심리학적 적절성을 평가하였다.

이러한 자신들의 연구나 작업 과정에서 언어학적 분석에 호소한 심리학자들(Abrahamsen, 1988; Gleitman과 Gleitman, 1970; Kaplan과 Bresnan, 1982; Klima와 Bellugi, 1979)의 또 다른 학문분야간 노력은, 단지 한 영역의 이론을 다른 영역에 적용하기만 하는 것을 넘어서서 다른 분야의 연구를 안내하기 위해 한 분야로부터 나온 이론적 관점 또는 결과를 재형식화하기도 한다. 그 과정에서, 심리학자들은 언어 처리와 언어의 추상적 구조간의 연결을 명세화할 수도 있으며, 그리하여 분야간을 포괄하는 분야 연계 이론을 산출할 수도 있다. 이 이론들은 심리학에서 연구된 기제가 어떻게 문화적 산물로서의 언어의 구조를 야기할 수 있는가를 보여준다. 그러나 이러한 탐구를 수행하는데 있어서 심리학자들이 언어학자들의 노력을 가로채려 시도하는 것도 아니며 심리학의 주도권을 주창하려는 것도 아니다.

심리학과 언어학간의 관계는 단방향적인 것이 아니며 언어학의 기술을 차용하여 수정함으로써 처리 기제에 대한 탐구를 안내하려는 심리학자들에 국한되는 것도 아니다. 심리학은 언어의 토대와 인간에 의해 사용될 수 있는 종류의 언어들에 대한 한계를 보여줌으로써 언어 현상을 설명하는 데 도움이 될 수 있다. 예를 들자면, 처리의 한계에 관한 심리학적 정보는 왜 특정 유형의 언어 구조가 더 큰 세력을 점유하고 있는지 설명해 줄 수도 있다. 이러한 현상은 심리학적 탐구를 통해서 설명될 수 있으나, 언어학 자체만으로는 설명될 수 없다. 그러므로 언어학 이론을 심리학적으로 사용하는 것과 언어학적 현상에 대해 심리학적 설명을 하는 것 양자 모두 학문분야의 경계를 넘나드는 이론적 개발로 귀결된다. 언어의 추상적 구조가 갖는 특징들은 언어를 이해하고 생성하는 체계의 특징과 연결된다.

### 3. 경계에 다리놓기 사례로서의 개념 및 범주화 연구

경계에 다리놓기 관계는 이론을 차용하거나 강제하는 것이 강조되지 않고 다른 학문분야들에서 각기 다른 지향점을 가지고 연구된 현



상들간의 연결을 끌어내는 것이 강조된다. 이런 경우, 각 학문분야의 연구자들은 자신이 관심을 가지는 현상들이 다른 현상들에 어떻게 관련되는가를 발견함으로써 관심을 가지는 현상들에 대한 그들 자신의 이해를 더욱 풍부히 할 수 있다.

인지과학에 참여하고 있는 다양한 학문분야의 종사자들이 궁극적으로 그 학문분야들간의 간극을 메워줄 수 있는 이론적 전망의 개발에 기여하고 있는 하나의 기획(project)은 개념과 범주화의 본성에 관한 연구이다. 개념은 여기서 사물의 범주에 대한 언어적 또는 정신적 표상으로 취급된다. 이러한 노력의 뿌리는 종종 철학자 Wittgenstein(1953)으로 거슬러 올라가는데, 그는 개념이 예컨대 그것에 상응하는 범주 속에 있기 위한 필요 충분 조건을 제공함으로써 정의될 수 있다는 고전적인 견해에 도전하였다. 그는 “게임”과 같은 간단한 개념에서조차도 필요 충분 조건을 진술하는 것은 불가능하다고 주장하였다. 그는 모든 게임들이 공유하는 공통된 속성의 집합이 있다기보다는, 어떤 주어진 게임이 그것의 전부가 아니라 일부만을 공유하는, 속성들의 중복되는(overlapping) 집합이 있을 것이라 제안하였다. 그는 그 상황을 사람의 가족 구성원간에 발견되는 특성들에 비유하였는데, 그들 각각은 특징들의 중복된 집합의 일부를 공유하지만, 어떤 특징도 가족 모두 그리고 오로지 그 가족들에 의해서만 공유되지 않는다.

개념의 고전적인 개념화에 대한 도전은 언어 구조에 관심이 있었던 두 인류학자 Berlin과 Kay(1969)에 의해 더 발전되었다. 서로 다른 문화들은 색깔을 단순히 다르게 분류한다고 일반적으로 믿어져 왔지만, 그들은 매우 다른 언어를 가진 서로 다른 문화들 내에서, 만일 그 문화가 어떤 특정 색에 대한 단어를 가지고 있다면, 사람들은 문화의 차이에도 불구하고 그 색의 중심적(focal)인 보기로 동일한 사례를 꼽는다는 것을 보여주었다. 사용하고 있는 색 용어의 적용범위에 있어서는 언어가 다를 수 있지만, 중심적 색깔로 여겨지는 것에 있어선 다르지 않다는 것이다. 이 작업은 심리학자 Eleanor Rosch와 그의 동료들에 의해 더 발전되었다. 초기에 Rosch는 기본 색깔들에

대한 용어를 결여하고 있는 언어를 가진 문화들에 있어서조차도 사람들은 그 색깔들의 중심적 사례들을 비중심적인 사례들보다 더 기초적인 방식으로 처리한다는 것을 보여주었다(Heider[Rosch], 1971, 1972).

Rosch는 그 다음으로 그녀의 생각을 색을 넘어 광범위한 개념들(예컨대 “새”와 같은)에 까지 확장하였고, 개념들 또한 유사한 구조를 지니고 있어서 한 범주의 어떤 사례들은(예컨대 참새) 다른 사례(예컨대 오리)들보다 더욱 전형적인(prototypical) 것으로 판단된다는 것을 보여주었다(Rosch, 1975; Rosch & Mervis, 1975). 언어적 증거에 부가하여 Rosch와 그녀를 이은 심리학자들은 개념의 구조에 대한 이러한 설명을 더 명료하게 하는 광범위한 심리학적 척도(예를 들면, 한 범주에서 보다 원형적인 사례를 범주화할 때는 더 빠른 반응시간을 보여주는 반응시간 척도)를 개발해 왔다.

Rosch는 그 후, 실제로 사람들이 한 대상이 어떤 범주에 속하는지에 대한 판단을 그것의 고유한 원형(prototype)에 대한 유사성을 기반으로 내린다는 것, 또는 개념들이 원형의 형태로 머리 속에 정신적으로 표상되어 있다는 것을 그녀가 수집한 증거들이 보여준다는 생각을 거부하였다. 그러나 수많은 심리학자들은 그러한 방식으로 그 결과를 해석한다. 예를 들어 Brooks(1978), 그리고 Medin과 Schaffer(1978)는 정신 표상의 범례 모형(exemplar model)을 주창하는데, 그 모형에 따르면 어떤 범주에 속한 개별 사례(어떤 이의 친숙한 개)가 범주 판단에 있어 준거점(reference point) 역할을 하며, 사람들은 다른 대상이 그 범주의 구성원인지를 그 사례에 대한 유사도의 관점에서 판단한다.

원형이라는 관점에서의 개념과 범주화에 대한 이러한 새로운 접근은 다양한 비판을 그만큼이나 다양한 학문분야로부터 끌어들었다. 심리학자의 두 집단, Osherson과 Smith(1981), 그리고 Armstrong, Gleitman과 Gleitman(1983)은, 원형 관점의 옹호자가 주장해 온 것처럼 범주가 글자 그대로 원형 구조를 가지고 있다는 견해에 대해 도전하였다. 예를 들어, Armstrong 등은 필요 충분 조건을 갖춘 개념의

전형적인 사례인 것으로 보이는 “홀수”와 같은 개념에 있어서조차도 원형 효과가 나타난다는 것을 보인다. Osherson과 Smith는 개념은 정의적 핵심(defining core: 범주의 구성원이 되기 위해서 만족시켜야 할 필요충분조건)과 식별 절차(원형성 결과를 유도하는)를 모두 포함한다는 견해를 옹호하였다. 철학자인 Georges Rey(1983) 역시 Rosch의 전통을 비판하는 데 동참하여, 언어적 의미를 정확하게 이해하는 것은 형이상학적 문제로서, 심리학적 사실로부터 분리하여 다루어져야 한다고 주장하였다. Putnam(1975b)의 뒤를 이어 그는 의미의 핵심은, 단어가 외부 세계와 연결되는 지시 관계에 놓여 있다고 여긴다. 지시된 대상들이 단어의 외연을 구성하고, 용어의 외연은 어떤 개인이 용어의 외연에 해당되는 것이 무엇인지를 결정하는 방식과는 독립적인 것으로 객관적인 것이다. 그렇게 하여 그는 원형성 판단이 언어의 의미에 어떠한 관계를 가진다는 주장도 거부하였다.

다른 이들은 다른 방식으로 비판에 동참하였다. 언어학자인 Anna Wierzbicka(1987)은 고전적인 견해(개념들은 범주의 구성원이 되기 위한 필요충분조건의 관점에서 정의된다는)의 옹호자들에 대한 Wittgenstein의 권고, 즉 범주의 구성원이 되기 위한 필요충분조건이 있다는 것을 단순히 전제하지만 말고 그러한 조건이 무엇인지를 보이라는 권고를 신중히 받아들였다. Wittgenstein이나 원형 개념을 채택해 온 학자들과는 달리, 그녀는 “게임”과 같은 개념에 대해 필요충분조건을 제공하는 것이 가능하다고 주장하였으며 정의의 후보가 될 만한 것들을 제시하였다. 그러나 그녀는 원형적 접근을 전적으로 폐기하지는 않았는데, 붉은, 새, 오르다와 같은 용어들의 정의는 그 필요충분조건 속에 비교적인 판단을 포함하므로, 정의 자체가 원형적인 사례에 의존한다고 논변하였다. 두 심리학자 Medin과 Ortony는, 이러한 다양한 비판들에 의지하여 심리학적 본질주의(psychological essentialism)를 옹호했는데, 그것에 따르면 개념의 표상은 본질을 부호화한다는 것이다. 그러나 그들은 이러한 본질이 표면상의 특성이 아니라 심층 이론들 속에 있는 것이라 여겼다. 그들은 이러한 심층

이론들이, 범주의 사례들을 식별하는 데 사용될 수 있는 보다 피상적인 면모들의 집합(예컨대, 시각적 면모들)을 산출하며, 우리의 시각 체계는 그러한 면모들에 잘 반응하도록 진화해 왔다고 주장하였다. 그러나 이 면모들은 정의하는 특성들이 아니다.

Rosch의 전통을 지지하는 수많은 이론가들은 원형성 판단을 설명하기 위해 이론과 같은 더 깊은 인지 구조를 끌어들이기 시작하였다. 심리학자인 Neisser(1986)는 범주화를 하는 우리의 능력이 세상 속 유기체의 패턴에 반응하는 시각 체계의 능력과 함께 시작할지 모르지만, 궁극적으로는 무엇이 공통의 범주에 속하는지 결정하는 이론에 의존한다고 주장하였다(Murphy와 Medin, 1985). 언어학자 George Lakoff(1987)는 “이상화된 인지모형”의 개념을 도입하여 우리가 대상을 범주화하는 데 영향을 미치는 지식을 조직화하는 단위를 지칭하였다. 이러한 모형은 우리가 사고하는 원리를 조직하고, 원형성 판단을 일으키는 것이라 주장하였다. 그러한 조직화하는 개념 중의 하나는 Lakoff가 철학자 Mark Johnson과 함께 비유에 대해 합동으로 연구하는 과정에서 개발한(Lakoff & Johnson, 1980) 환유 모형(metonymic model)의 개념이다. 이것은 추론과 재인의 작업에서 범주의 한 부분이 범주 전체를 대표한다는 모형이다. 예를 들자면 어머니의 개념 중 하나인 가정주부인 어머니가 전체 범주를 대표하는 일을 한다는 것이다. 이러한 일이 발생하면, 직장인 어머니는 가정주부인 어머니보다 덜 좋은 사례라는 원형성 판단이 초래된다. 또 하나의 조직화 원리는 중심적인 사례에서부터 출발하여 변화하는 다양한 사례들을 정의할 때 발생하는, 방사형 범주를 수반한다. ‘어머니’의 예를 다시 들어보면, 우리는 아이를 낳고, 그 아이의 유전자의 반을 제공하며, 아이를 기르는 어머니 등의 중심적인 사례로부터 시작하여 이 패턴으로부터 벗어나는 특수한 사례들을 정의할 수 있다(예를 들자면 계모, 양모, 생모). 이러한 모형들은 모두 우리의 이상적인 모형으로부터 벗어나며, 또한 Lakoff에 따르면 원형성 판단을 산출하는 데 기여한다. Lakoff는 개념과 범주는 이러한 심층적 인지모형의 산물이며,

그러므로 단순한 사과의 원자(atom)가 아니라고 주장하였다.

이러한 개념과 범주화에 관한 작업에서, 우리는 인지과학의 서로 다른 분야로부터 같은 주제, 즉 우리가 가지고 있는 개념의 본성에 관해 언급하는 수많은 연구자들을 발견한다. 그러나 서로 다른 분야에서 온 결과로 인해, 그들은 서로 다른 이론적 모형들과 서로 다른 연구의 양식을 개념과 범주화에 관한 작업에 도입하였다. 원형성 판단에 관한 초기의 결과에 관하여 상당한 불일치가 있었음에도 불구하고, 연구자들이 작업을 해 가는 방향은 인지과학에 포함된 서로 다른 학문분야의 요구들에 적절한 개념과 범주화에 대한 통합된 설명으로 향하고 있는 것이다. 그러한 설명은 한 학문 분야를 다른 것으로 어떻게든 환원함이 없이 학문분야간에 다리를 놓는 분야 연계 이론의 또 다른 사례를 제공하고 있다.

## V. 포괄적 인지과학 방법론으로서의 분야간 이론(interfield theory)

지금까지 우리는 인지과학 방법론으로 수준간 방법론과 수준내 방법론을 살펴보았다. 수준간을 연계시키는 방법론은 상위수준에 위치한 인지과학과 하위수준에 위치한 신경과학을 연결시키는 방법론으로 '환원', '기능적 분석', '병렬분산처리' 등을 그 예로서 고찰하였다. 그리고 동일 수준에 속하는 학문분야들을 연계시키는 방법론은 심리학, 인공지능, 언어학, 인류학, 그리고 철학 등과 같이 동일하게 상위수준에 속하는 인지과학내의 학문 분야들을 연결시키는 방법론으로 '경계 허물기'와 '경계에 다리놓기'를 그 예로서 고찰하였다.

이러한 여러 방법론을 모두 아우를 수 있는 보다 포괄적인 인지과학 방법론은 없을까? 그것에 대한 답은 '포괄적'이라는 용어의 의미를 얼마나 엄밀하게 받아들이느냐에 따라 달라질 수 있다는 것이 필자의 생각이다. 다시 말해, 위의 학문 분야의 이론들을 모두 하나의

들에 맞춰 통일적으로 설명해낼 수 있는 그러한 통일적인 방법론적 틀을 요구한다면, 그러한 방법론의 수립은 어렵다는 것이다. 그러나 다른 학문분야들에서 각기 다른 지향점을 가지고 연구된 현상들간의 연결을 끌어내고 그들간의 관계에 대한 전체적인 조망을 가능케 하는 것으로서의 포괄적 방법론이라고 한다면, 그것은 가능하다는 것이다. 필자는 후자의 방법론에 대한 이름으로 '분야간 이론(interfield theory)'이란 용어를 Darden과 Maull 그리고 Bechtel을 좇아 제안한다.

Bechtel은 '분야간 이론'을 '환원'과 대비되는 용어로 사용했다는 점에서, 필자의 용어법과는 다르다. 필자의 생각으로, 어차피 학문 분야들간의 전반적인 환원이 어려운 상황이라면 (그것이 비록 잠정적으로 어려운 것이라 할지라도), 부분적인 환원도 전체적인 관계 설정에 대한 이해를 깊게 해준다는 의미에서 분야간 이론의 일종으로 볼 수 있다는 것이다. 따라서 필자의 포괄적 인지과학 방법론으로서의 '분야간 이론'은 위에서 말한 '환원', '기능적 분석', '병렬분산처리', '경계 허물기', '경계에 다리놓기' 등을 모두 포괄하는 방법론이다. 그것들의 공통점은 다른 학문분야들에서 각기 다른 지향점을 가지고 연구된 현상들간의 연결을 끌어내고 그들간의 관계에 대한 전체적인 조망을 가능케 해준다는 점에 있다. 비유하여 말하자면, 인지과학 방법론은 적어도 지금 시점에서는 모든 다양한 종류의 철광석들을 하나의 방식으로 녹여내는 용광로(melting pot)일 수는 없으며, 다양한 재료들을 모아 그 전체성에서 조화를 이루어내는 샐러드 보울(salad bowl)이 될 수밖에 없다는 것이 필자의 결론이다.

연구 방법론은 연구 분야의 내용과 완전히 독립적으로 결정될 수는 없다. 더구나 경험적 사태에 닿을 내리고 있는 인지 현상에 대한 연구는 선형적 방법만으로 또는 단 한가지 방식의 경험적 방법만으로 모든 연구를 마무리지을 수는 없을 것이다. 인지과학 연구는 다원적 형태를 아우르는 방식의 방법론으로 진행되어야 할 것이다.

## 참고문헌

- 김영정(1996), 『심리철학과 인지과학』, 철학과 현실사
- 조인래(2001), “인지과학의 방법: 기능적 분석”, 『철학사상』 제12호, 189-219쪽, 서울대학교 철학사상연구소.
- Abrahamsen, A. A.(1987), Bridging boundaries versus breaking boundaries: Psycholinguistics in perspective. *Synthese*, 72, 355-388.
- \_\_\_\_\_ (1988), *Bridging interdisciplinary boundaries: The case of kin terms.*
- Armstrong, S. L., Gleitman, L., & Gleitman, H.(1983), What some concepts might not be. *Cognition*, 13, 263-308.
- Bechtel, W.(1985), Teleological functional analyses and the hierarchical organization of nature. In N. Rescher (Ed.), *Teleology and natural science*(pp.26-48) Landham, MD: University Press of America.
- \_\_\_\_\_ (1986), What happens to accounts of the mind-brain if we forego an architecture of rules and representations? In A. Fine & P. Machamer (Eds.), *PSA 1986*(Vol. 1, pp.159-171). East Lansing, MI: Philosophy of Science Association.
- \_\_\_\_\_ (1988a), *Philosophy of Mind: An overview for cognitive science.* Hillsdale, NJ: Lawrence Erlbaum Associates.
- \_\_\_\_\_ (1988b), *Philosophy of Science: An overview for cognitive science.* Hillsdale, NJ: Lawrence Erlbaum Associates.

- Bechtel, W., & Richardson, R. C.(1988), *A model of theory development: Localization as a strategy in scientific research*.
- Berlin, B., & Kay, P.(1969), *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.
- Blumenthal, A. L.(1987), The emergence of psycholinguistics. *Synthese*, 72, 313-324.
- Boveri, T.(1903), Über die Konstitution der chromatischen Kernsubstanz. *Verhandlungen der deutschen zoologischen gesellschaft zu Würzburg*, 13, 10-13.
- Bradley, D., Garrett, M., & Zurif, E.(1980), Syntactic defects in Broca's aphasia. In D. Caplan (Ed.), *Biological Studies of mental processes*(pp.269-286). Cambridge, MA: MIT Press.
- Broca, P.(1861), Remarques sur le siege de la faculte du langage articule, suivies d'une observation d'aphemie. *Bulletin de la Societe Anatomique de Paris*, 6, 343-357.
- Brooks, L.(1978), Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*(pp.169-211). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Chomsky, N.(1968), *Language and mind*. New York: Harcourt, Brace, & World.
- Churchland, P. S.(1986), *Neurophilosophy: Toward a unified science of the mind-brain*. Cambridge, MA: MIT Press/Bradford Books.
- Clark, H. H., & Clark, E. V.(1977), *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt, Brace, Jovanovich.



- Crick, F. H. C.(1984), Function of the thalamic reticular complex: The searchlight hypothesis. *National Academy of Sciences of the United States of America. Proceedings. Biological Sciences*, 81, 4586-4590.
- Cummins, R.(1975), Functional Analysis, *Journal of Philosophy*, 72, 741-765.
- \_\_\_\_\_ (1983), *The Nature of Psychological Explanation*, MIT Press.
- Darden, L., & Maull, N.(1977), Interfield theories. *Philosophy of Science*, 43, 44-64.
- Denett, D. C.(1981), *Brainstorms*. Cambridge, MA: MIT Press/Bradford Books.
- Fodor, J. A.(1974), Special Sciences(Or: Disunity of science as working hypothesis). *Synthese*, 28, 97-115.
- \_\_\_\_\_ (1978), Propositional attitudes. *The Monist*, 61, 501-523.
- Gall, F. J.(1835), *Sur les fonctions due cerveau et sur celles de chacune de ses parties*(W. Lewis, Trans.). (Originally published, 1809)
- Geschwind, N., & Kaplan, E.(1962), A human cerebral disconnection syndrome. *Neurology*, 12, 675-685.
- Gleitman, L. R., & Gleitman, H.(1970), *Phrase and paraphrase: Some innovative uses of language*. New York: Norton.
- Grodzinsky, Y., Swinney, D., & Zurif, E.(1985), Agrammatism: Structural deficits and antecedent processing disruptions. In M. L. Kean (Ed.), *Agrammatism*(pp.65-81). New York: Academic Press.
- Heider [Rosch], E.(1971), Focal' color areas and the development of color names. *Developmental Psychology*, 4, 447-455.
- \_\_\_\_\_ (1972), Universals in color naming and memory.

- Journal of experimental psychology*, 93, 10-20.
- Kaplan, R. M., & Bresnan, J.(1982), Lexical-functional grammar: A formal system for grammatical representation. In J. Bresnan (Ed.), *The mental representation of grammatical relations*(pp.173-281). Cambridge, MA: MIT press.
- Klima, E., & Bellugi, U.(1979), *The signs of language*. Cambridge, MA: Harvard University Press.
- Lakoff, G.(1987), *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Lakoff, G., & Johnson, M.(1980), *Metaphors we live by*. Chicago: University of Chicago Press.
- Liebig, J.(1842), *Animal chemistry or organic chemistry in its application to physiology and pathology*(W. Gregory, Trans.). Cambridge, England: John Owen.
- Lycan, W. G.(1981), Form, function, and feel. *The Journal of Philosophy*, 78, 24-49.
- Machamer, P.(1977), Teleology and selection processes. In R. G. Colodny (Ed.), *Logic, laws, and life: Some philosophical complications*(pp.129-142). Pittsburgh: University of Pittsburgh Press.
- McCauley, R. N.(1987), The not so happy story of the marriage of linguistics and psychology, or why linguistics has discouraged psychology's recent advances. *Synthese*, 72, 341-354.
- McClelland, J. L., Rumelhart, D. E., & the PDP Research Group. (1986), *Parallel distributed processing: Explorations if the microstructures of cognition, Vol. 2: Psychological models*. Cambridge, MA: MIT press/Bradford Books.

- Medin, D. L., & Schaffer, M. M.(1978), Context theory of classification learning. *Psychological Review*, 85, 207-238.
- Murphy, G. L., & Medin, D. L.(1985), The role of theories in conceptual coherence, *Psychological Review*, 92, 289-316.
- Neisser, U.(1986), From direct perception to conceptual structure. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization*. Cambridge, MA: Cambridge University Press.
- O'Keefe, J., & Nadel, L.(1978), *The hippocampus as a cognitive map*. Oxford: Clarendon Press.
- Osherson, D., & Smith, E. E.(1981), On the adequacy of prototype theory as a theory of concepts. *Cognition*, 9, 35-58.
- Pellionisz, A., & Llinas, R.(1982), Space-time representation in the brain. The cerebellum as a predictive space-time metric tensor. *Neuroscience*, 7, 2249-2970.
- \_\_\_\_\_ (1985), Cerebellar function and the adaptive feature of the central nervous system. In A. Berthoz & G. Melvill Jones (Eds.), *Adaptive mechanisms in gaze control*(pp.223-232). Amsterdam: Elsevier
- Putnam(1975a), Philosophy and our mental life. In H. Putnam (Ed.), *Mind, language, and reality: Philosophical papers of Hilary Putnam*(Vol. 2, pp.291-303). Cambridge, MA: Cambridge University Press.
- \_\_\_\_\_ (1975b), The meaning of 'meaning'. In H. Putnam (Ed.), *Mind, language, and reality: Philosophical papers of Hilary Putnam*(Vol. 2, pp.215-271). Cambridge, MA: Cambridge University Press.
- Reber, A.(1987), The rise and (surprisingly rapid) fall of psycholinguistics. *Synthese*, 72, 325-340.

- Rey, G.(1983), Concepts and stereotypes. *Cognition*, 15, 237-262.
- Richardson, R. C.(1979), Functionalism and reductionism. *Philosophy of Science*, 46, 533-558.
- Rosch, E.(1975), Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.
- Rosch, e., & Mervis, C.(1975), Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986), *Parallel distributed processing. Explorations in the microstructures of cognition. Vol. 1: Foundations.* Cambridge, MA: MIT Press/Bradford Books.
- Rumelhart, D. E.(1984), The emergence of cognitive phenomena from the sub-symbolic processes. *Proceedings of the sixth annual conference of the Cognitive Science Society* (pp.59-62). Boulder, CO.
- Schaffner, K.(1967), Approaches to reduction. *Philosophy of Science*, 34, 137-147.
- Simon, H. A.(1980), *The sciences of the artificial* (2nd ed). Cambridge, MA: MIT Press.
- Sutton, W.(1903), The chromosomes in heredity. *Biological Bulletin*, 4, 231-251.
- Von Eckardt, B.(1978), Inferring functional localization from neurological evidence. in E. Walker(Ed.), *Explorations in the biology of language*(pp.27-66). Cambridge. MA: MIT Press/Bradford Books.
- Wierzbicka, A.(1987), 'Prototypes save': On the current uses and abuses of the concept 'prototype' in current linguistics,

*philosophy, and psychology*. Unpublished manuscript.

Wimsatt, W. C.(1972), Teleology and the logical structure of function statements. *Studies in the History and Philosophy of Science*, 3, 1-80.

\_\_\_\_\_ (1976a), Reductionism, levels of organization, and mind-body problem. In G. Globus, G. Maxwell, & I. Savodnik (Eds.), *Consciousness and the brain: A scientific and philosophical inquiry*(pp.205-267). New York: Plenum Press.

\_\_\_\_\_ (1976b), Reductive explanation: A functional account. In R. S. Cohen, C. A. Hooker, A. C. Michalos, & J. van Evra (Eds), *PSA-1974, Boston studies in the philosophy of science*(Vol. 32, pp.671-710).

\_\_\_\_\_ (1979), Reduction and reductionism. In P. D. Asquith & H. Kyburg (Eds.), *Current research in philosophy of science*(pp.352-377). East Lansing: Philosophy of Science Association.

Wittgenstein, L.(1953), *Philosophical investigations*. New York: Macmillan.

Wright, L.(1973), Functions, *Philosophical Review*, 82, 139-168.

\_\_\_\_\_ (1976), *Teleological Explanations*, Berkeley: University of California Press.